DOCUMENT RESUME

ED 440 550 FL 026 249

TITLE	JALT Journal, 1997.
INSTITUTION	Japan Association for Language Teaching, Tokyo.
ISSN	ISSN-0287-2420
PUB DATE	1997-00-00
NOTE	364p.
AVAILABLE FROM	JALT Central Office, Urban Edge Bldg. 5F, 1-37-9 Taito, Taito-Ku. Tokyo 110 Japan. Tel: 81-3-3837-1630.
PUB TYPE	Collected Works - Serials (022)
LANGUAGE	English, Japanese
JOURNAL CIT	JALT Journal; v19 n1-2 May-Nov 1997
EDRS PRICE	MF01/PC15 Plus Postage.
DESCRIPTORS	Advertising; Code Switching (Language); Contrastive Linguistics; Critical Thinking; *English (Second Language); Essay Tests; Foreign Countries; Interlanguage; *Japanese; Poetry; Pronunciation; Second Language Instruction; Second Language Learning; Student Evaluation; Teaching Methods
IDENTIFIERS	Australia; Japan

ABSTRACT

The 1997 issues of "JALT Journal" include the following articles: "Influence of Learning Context on Learners' Use of Communication Strategies"; "The Eiken Test: An Investigation"; "Assessing EFL Student Progress in Critical Thinking with the Ennis-Weir Critical Thinking Essay Test"; "Contrastive Rhetoric in Letter Writing"; "Japanese EFL Learners' Test-Type Related Interlanguage Variability"; "Codeswitching in EFL Learner Discourse"; "Japanese and American Television Commercials: A Cultural Study with TEFL Applications"; "Validating a Questionnaire on Confidence in Speaking English as a Foreign Language"; "On Reading-Writing Relationships in First and Foreign Languages"; "Teaching with Music: A Comparison of Conventional Listening Exercises with Pop Song Gap-Fill Exercises"; "Learning Pronunciation and Intonation of Japanese through Drama by Beginning Language Students: A Case for Reflective Journals"; "Course Design and Delivery of Hospitality Japanese in Australia" (in Japanese); "Empathy and Teacher Development"; "Answer, Please Answer! A Perspective on Japanese University Students' Silent Response to Questions"; "A Poem in the Process: Haiku as an Alternative to Brainstorming." Issues include articles, research forum, perspectives, and review sections. (Contains extensive references.) (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

JALT Journal

Volume 19, Numbers 1-2

May-November 1997

ED 440 550

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

G. Troyer

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL 026 249

BEST COPY AVAILABLE

SN 0287-2420 ¥950

JALT JOURNAL

全国語学教育学会

May 1997

Volume 19 • No. 1

Japan Association for Language Teaching

Inside this issue:

Communicative strategies • *Eiken* Test
Critical thinking • Letter writing
Performance activities
Test-type variability • Codeswitching
Television Commercials

Your Ticket to the World!

The Diploma Program in Teaching English as a Second or Foreign Language (TESL/TEFL) at the Saint Michael's School of International Studies

Vermont is well known as a US travel destination, but for those who add the Diploma Program to their itinerary, it is also a gateway to international teaching opportunities. The TESL/TEFL Diploma Program is an intensive eight-week summer program for native and non-native speakers of English. It has a practical focus, and may help graduates find TESL or TEFL positions abroad.

Applications are now being accepted for the June 16th to August 8th session.

A LEADER IN ESL/EFL EDUCATION FOR OVER 40 YEARS

School of International Studies
Saint Michael's University
Winooski Park, Colchester, Vermont 05439-0253 USA
Phone: (802) 654-2300 • Fax: (802) 654-2595
E-mail: sis@smcvt.edu • Website: http://www.smcvt.edu

Also at the School of International Studies—

Teacher Education Programs
Master of Arts in TESL/TEFL
Advanced Certificate in TESL/TEFL
Institute in TESL/TEFL

Launch Your Career

in the Teaching of ESL with a Master's in TESL from Saint Michael's!

Our Master of Arts in Teaching English as a Second (or Foreign) Language gives you the knowledge and skills you need for a successful teaching career.

- Complete our 36-credit program in one year of full-time study, part time, or in three summers.
- Or select the 39-credit with thesis option.
- Receive a high level of individual attention and career counseling.
- Gain useful contacts for finding an excellent job.
- Certificate programs of 9 and 18/21 credits also available.

School of International Studies / Saint Michael's University
Winooski Park, Colchester, Vermont 05439-0253 USA
Phone: (802) 654-2300 • Fax: (802) 654-2595 • E-mail: sis@smcvt.edu
Website: http://www.smcvt.edu

Contents

May 1997
Volume 19 • No. 1

4 In this issue
6 From the Editors

Articles

7 Influence of Learning Context on Learners' Use of Communication Strategies
Ryu Kitajima
24 The Eiken Test: An Investigation
Laura MacGregor
43 Assessing EFL Student Progress in Critical Thinking with the Ennis-Weir Critical Thinking Essay Test
Bruce W. Davidson & Roaney A. Dunham
58 Contrastive Rhetoric in Letter Writing
Taeko Kamimura & Kyoko Oi
77 実際使用アクティビティーを使用した海外での日本語教育例
加藤好崇

Research Forum

89 Japanese EFL Learners' Test-Type Related Interlanguage Variability
Akihiro Ito
106 Codeswitching in EFL Learner Discourse
Ethel Ogane

Perspectives

123 Japanese and American Television Commercials: A Cultural Study with TEFL Applications
Paul Stapleton

Reviews

138 *Multilingual Japan* (John C. Maher & Kyoko Yashiro, Eds.) – Steve McCarthy
142 *Oxford Advanced Learner's Dictionary* – Brian C. Perry
147 *Understanding Communication in Second Language Classrooms* (Karen E. Johnson) – Damian Lucantonio
151 *An Introduction to Spoken Interaction* (Anna-Brita Stenström) – Hugh Molloy
156 *A Communicative Grammar of English* (Geoffrey Leech & Jan Svartvik) – Kevin Varden
161 *Functional English Grammar: An Introduction for Second Language Teachers* (Graham Lock) – Randall Davis
163 *Spoken English on Computer: Transcription, Mark-up and Application* (Geoffrey Leech, Greg Myers, & Jenny Thomas, Eds.) – Charles Adamson
165 Books to Review

JALT Journal Information

166 Information for Contributors (English and Japanese)

All materials in this publication are copyright ©1997 by their respective authors.

Japan Association for Language Teaching

JALT is a professional organization dedicated to the improvement of language learning and teaching in Japan, a vehicle for the exchange of new ideas and techniques, and a means of keeping abreast of new developments in a rapidly changing field. Formed in 1976, JALT has an international membership of more than 4000. There are currently 38 JALT chapters throughout Japan. It is the Japan affiliate of International TESOL (Teachers of English to Speakers of Other Languages) and a branch of IATEFL (International Association of Teachers of English as a Foreign Language).

JALT publishes *JALT Journal*, *The Language Teacher* (a monthly magazine of articles and announcements on professional concerns), *JALT Applied Materials* (a monograph series), and JALT International Conference proceedings.

The JALT International Conference on Language Teacher/Learning and Educational Materials Exposition attracts some 2000 participants annually. Local meetings are held by each JALT chapter, and National Special Interest Groups disseminate information on specific concerns. JALT also sponsors specials events, such as conferences on specific themes.

JALT provides awards for Research Grants and Development, announced annually at the conference.

Membership, open to those interested in language education, includes enrollment in the nearest chapter, copies of JALT publications, and reduced admission to JALT sponsored events. For information on membership, contact the JALT Central Office.

JALT National Officers (1997)

President: Gene van Troyer

Vice President: Brendan Lyons
Treasurer: Larry J. Cisar
Recording Secretary: Jim Chambers

Programs: Carolyn Latham
Membership: Richard Marshall
Public Relations: Jane Hoelker

Chapters

Akita, Chiba, East Hokkaido (affiliate), Fukui, Fukuoka, Gunma, Hamamatsu, Himeji, Hiroshima, Hokkaido, Ibaraki, Iwate, Kagawa, Kagoshima, Kanazawa, Kitakyushu (affiliate), Kobe, Kochi, Kumamoto (forming) Kyoto, Matsuyama, Miyazaki (affiliate), Nagano, Nagasaki, Nagoya, Nara, Niigata, Okayama, Okinawa, Omiya, Osaka, Sendai, Shizuoka, Suwa, Tochigi, Tokushima, Tokyo, Toyohashi, West Tokyo, Yamagata, Yamaguchi, Yokohama

National Special Interest Groups

Bilingualism; Computer Assisted Language Learning; College and University Educators; Global Issues in Language Education; Japanese as a Second Language; Junior and Senior High School Teaching; Learner Development; Material Writers; Other Language Educators; Professionalism, Administration and Leadership in Education; Second Language Literacy; Teacher Education; Teaching Children; Testing and Evaluation; Video

JALT Central Office

Junko Fujio, Office Manager
Urban Edge Building, 5th Floor
1-37-9 Taito, Taito-ku, Tokyo 110 Japan

JALT Journal

Volume 19, No. 1
May, 1997

Editor
Tamara Swenson
Osaka Jogakuin Junior College

Reviews Editor
Thomas Hardy
Tamagawa University

Associate Editor
Sandra Fotos
Senshu University

Japanese-Language Editor
Naoko Aoki
Shizuoka University

Editorial Board

Charles E. Adamson, Jr.
Miyagi University
Fred Anderson
Fukuoka University of Education
William Bradley
Ryukoku University
James Dean Brown
University of Hawaii, Manoa
Christine Pearson Casanave
Keio University, SFC
Roger Davies
Ehime University
Ann Ediger
Teachers College, Columbia University
Rod Ellis
Temple University
John Fanselow
Teachers College, Columbia University
Kevin Gregg
Momoyama University
Dale T. Griffee
Segakuin University
Paul Gruba
University of Melbourne
Ilona Leki
University of Tennessee
Mary Goebel Noguchi
Ritsumeikan University
David Nunan
City University of Hong Kong
Akito Ozaki
Nagoya University
Tim Riney
International Christian University
Carol Rinnert
Hiroshima City University
Thomas Robb
Kyoto Sangyo University
Patrick R. Rosenkjar
Temple University Japan
Tadashi Sakamoto
Nanzan University
Bernard Susser
Doshisha Women's Junior College
Satoshi Toki
Osaka University
Deryn Verity
Osaka University

Additional Readers: *William Acton, Michael Higgins, Nöel Houck, Guy Modica, Brad Visgatis*
JALT Journal Proofreading: Greta Gorsuch, Takuo Kinugawa, Brad Visgatis, Jack Yohay

JALT Publications Board Chair: William Acton

The Language Teacher **Editor**: Antony Cominos **Associate Editor**: Steve McGuire

Layout: The Word Works **Cover**: Amy Johnson **Printing**: Koshinsha

In this issue

Articles

Five articles are included in this issue. **Ryu Kitajima** examines the influence of a meaning-focused versus a form-focused learning context upon three groups of learners' choices of communication strategies and their overall communicative performance. Results indicate that subjects receiving meaning-focused instruction significantly reduced their use of reduction strategies and increased their use of achievement strategies.

Laura MacGregor undertakes an examination of the *Eiken* tests, administered by the *Eigo Kentei Kyokai* (*Eikyo*), or in English the Society for Testing English Proficiency (STEP). Her study seeks to further understand the *Eiken* tests through 1) an examination of *Eikyo's* statements about the tests and 2) administration of one pre-second level test to a group of Japanese college students and use of these results for studies of the test's reliability and validity.

The importance of promoting thinking has received considerable attention, however, research has not established that training in thinking skills can be effectively combined with EFL/ESL instruction. **Bruce W. Davidson** and **Rodney A. Dunham** report on the use of the Ennis-Weir Critical Thinking Essay Test to assess progress in critical thinking after academic instruction. Results indicate that critical thinking skills can be taught in the EFL classroom.

The relationship between Japanese college students' proficiency in English and their cultural awareness toward a target-language culture (American) are examined by **Taeko Kamimura** and **Kyoko Oi**. Results show that subjects with both high English proficiency and cultural awareness manifested the rhetorical patterns closest to those of native speakers, indicating that cultural awareness may be as important an element as English ability in student writing.

One article in Japanese is included in this issue. **Yoshitaka Kato**, examining the outcome of project work by advanced learners of Japanese at a Russian university, concludes from his study that learner interviews of NSs add to and reinforce the learners' vocabulary and understanding of dialects, and help learners become aware of the way Japanese view Russian customs.

Research Forum

Two articles are included. The first, by **Akihiro Ito**, investigates the effects of differences of test-types on the accuracy rates in interlanguage performance of Japanese EFL learners. The second, by **Ethel Ogane**, looks at codeswitching by adult Japanese EFL learners, suggesting that codeswitching with these learners is discourse related.

Perspectives

Paul Stapleton examines the cultural values of Japanese and American television commercials and suggests they can provide a rich source of both cultural information for teachers interested in promoting learner awareness of target culture values.

Reviews

Texts on multinlingualism in Japan, classroom language use, conversational analysis, grammar, and the use of computer-stored corpora, as well as one learner dictionary, are reviewed by **Steve McCarthy**, **Brian C. Damian Lucantonio**, **Hugh Molloy**, **Kevin Varden**, **Randall David**, and **Charles Adamson**.

From the Editors

JALT On-Line

JALT has expanded its presence on the World Wide Web with several web sites. JALT general information can be obtained at <http://langue.hyper.chubu.ac.jp/jalt/>. Information from JALT's monthly magazine, *The Language Teacher*, is available at <http://langue.hyper.chubu. ac.jp/jalt/pub/tlt/>.

Conference News

The **23rd Annual JALT International Conference** on Language Teaching/Learning and Educational Materials Exposition is scheduled for October 9 to 12, 1997 in Hamamatsu. The theme of the 1997 JALT Conference is Trends and Transition. Those interested in attending the conference are urged to contact the JALT Central Office for more information or visit the JALT '97 web page at <http://www.miyazaki-mic.ac.jp/jalt/JALT97.html>.

The **3rd Annual Pacific Second Language Research Forum**, scheduled for March 26 to 29, 1998, invites proposals for papers and reports of work in progress on data-based research into second language acquisition and use. Papers and reports may be given in either English or Japanese. The deadline for submission of proposals is September 1, 1997. Participants are being sought for symposia on: models of second language processing; language attrition; Japanese as a second language; fossilization; task design and interlanguage; focus on form in instructed SLA; bilingualism; child second language acquisition; psycholinguistic perspectives on the relationship between L2 reading and listening; SLA studies in generative grammar, and cognition and SLA. Proposals should be sent to: Peter Robinson, Aoyama Gakuin University, Dept. of English (PacSLRF '98), 4-4-25 Shibuya, Shibuya-ku, Tokyo 150 Japan. Information about PacSLRF '98 or proposal submission requirements can be obtained from the PacSLRF '98 web site <http://www.als.aoyama.ac.jp/pacslrf.html>.

Corrections

Editorial Advisory Board member **Anne Ediger**'s affiliation was incorrectly listed in *JALT Journal, 18*(2). Her correct affiliation is Teachers College, Columbia University.

Additional Reader **Stephen M. Ryan**'s name was incorrectly given in *JALT Journal, 18*(2).

We sincerely appreciate their efforts on JALT's behalf and wish to apologize for any inconvenience this may have caused.

Articles

Influence of Learning Context on Learners' Use of Communication Strategies[1]

Ryu Kitajima
San Diego State University

This study examines the influence of a meaning-focused versus a form-focused learning context upon learners' choices of communication strategies and their overall communicative performance. For this purpose, an 11-week study was conducted with three groups of students at a university in Japan. The control group studied English in a form-focused learning context. The experimental groups (1 and 2) studied English in a meaning-focused learning context. In addition, experimental group 2 received training in communication strategies. Two types of communicative tasks, 1) description of pictures and 2) narration of a picture story, were administered before and after treatments. Communication strategies were identified by two raters, based on the subjects' audio-taped communicative performance and immediate retrospection. Performances were evaluated separately by the same raters. The results show that both experimental groups significantly reduced use of reduction strategies and increased use of achievement strategies. Similarly, results show that communicative performance by the experimental groups was evaluated higher than that of the control group on the post-test. The findings suggest that learning context has an important influence on learners' use of communication strategies and their communicative performance.

本研究は、意味重視の学習と形式重視の学習が学習者のコミュニケーション法略の選択ならびに全般的コミュニケーションのパフォーマンスに与える影響を考察する。日本の大学において3つのグループが11週間にわたって調査された。統制群は形式重視の文脈で英語を学習し、二つの実験群は意味重視の文脈で学習した。実験群の一つはコミュニケーション法略のトレーニングも受けた。2種類のコミュニカティブなタスクによるテストが事前、事後に実施された。タスクの音声録音とタスク遂行直後の被験者の内省をもとに、コミュニケーション法略が二人の判定者により認定された。パフォーマンスの評価は別途行った。その結果、二つの実験群で、縮小法略が有意に減少し、到達法略の使用が増加した。さらに、事後テストにおいて実験群のパフォーマンスは統制群より高く評価された。これらの結果は、学習の文脈がコミュニケーション法略の使用と、コミュニケーションのパフォーマンスに重要な影響を与えることを示している。

JALT Journal, Vol. 19, No. 1, May, 1997

BEST COPY AVAILABLE

During the past two decades, second language (L2) research has focused on learners' strategies for learning and using a language. One such area involves studies of communication strategies (CSs) (Bialystok, 1990; Bongaerts, Kellerman & Bentlage, 1987; Bongaerts & Poulisse, 1989; Chen, 1990; Dornyei, 1995; Faerch & Kasper, 1983b; Paribakht, 1985; Poulisse & Schils, 1989; Tarone, 1983). CSs, which are learners' attempts to solve communicative problems occurring in the middle of realizing a certain meaning in spontaneous communication, are considered indispensable components of communicative performance. Since learners in communicative situations often encounter difficulty in realizing intended meaning because of limited linguistic resources, they resort to CSs in order to maintain communication.

CSs are generally defined either from an interactionist view (Tarone, 1983) or a psycholinguistic view (Faerch & Kasper, 1983a, 1983b). While the interactionist view emphasizes the learner and the interlocutor's mutual attempts to solve a communicative problem through negotiation of meaning, the psycholinguistic view narrows CSs to those within the learner's internal mechanism. It sets up two criteria for its operation: 1) the existence of a problem which the learner faces for actualizing intended meaning in the target language, caused by insufficient linguistic resources or by the difficulty of retrieving relevant linguistic items from the memory system, and 2) the learner's awareness of the existence of the problem and the necessity of solving it (Faerch & Kasper, 1983a). This study adopts the second definition.

CS studies have found that the use of some CSs results in more effective problem solving than others (Corder, 1983; Faerch & Kasper, 1983a; Tarone, 1983) and that the use of CSs contributes to variation in the overall effectiveness of learners' communicative performance (Chen, 1990). For example, reduction strategies such as *Message Abandonment* (Tarone, 1983) direct the learner to avoid solving a problem and to give up on conveying the message. On the other hand, achievement strategies such as *Analytic* and *Holistic* strategies (Poulisse, 1987) direct the learner to work on an alternative plan for reaching the original goal by means of whatever resources are available. These findings suggest that, in order to understand the variability in learners' communicative abilities in a language, we must study the factors which contribute to the differences in learners' uses of CSs.

Cognitive views of L2 learning (Bialystok, 1990; Bialystok & Sharwood-Smith, 1985; Ellis, 1986; O'Malley & Chamot, 1990; Tarone, 1983) hold that the ability to use language does not come about as a result of an increase in the static rules of the language. Rather, it requires cognitive strategies and processes which control the knowledge. These cognitive views consider the type of language task a major factor in the proceduralization of language knowledge, by requiring particular cognitive strategies for effective performance.

These views raise questions in terms of the role of learners' strategies in mediating a learning context and their learning outcome. Does a learning context which directs learners to engage in similar tasks for an extended period of time effect acquisition of particular strategies? Do learners in a particular learning context search, use, and acquire particular strategies to process information effectively and efficiently?

The Study

This study investigates whether learners exposed to meaning-focused learning contexts use CSs in communicative tasks differently from those in form-focused learning contexts.

Method

Subjects: An 11-week study was conducted with 15 students of English at a four-year national women's university in Japan. Of the 15 students, 13 were sophomores, one a junior, and one a freshman. Majors were home economics (four), humanities (three), biology (two), and architecture (one). The Ss had studied English for more than six years in required junior and senior high school classes. Though the Ss were taking an English course, as required by the university when they participated in this study, the courses available focused on reading translation, not on English for communication. All Ss voluntarily participated in the study to improve their listening and speaking skills. The Ss English proficiency was considered at least intermediate in terms of grammar in that they had achieved scores high enough for admission on the English portion of the standardized entrance exam administered by the Japanese Ministry of Education. However, all Ss demonstrated difficulty in expressing even simple meanings in spoken English on the pre-test (discussed below). Ss were randomly assigned to either the control group or one of the two experimental groups (Ex1 and Ex2), with five Ss in each group.

Treatment: The control group studied English in a form-focused instructional context. The experimental groups studied English in meaning-focused instructional contexts. In Ex1, CSs were allowed to evolve without explicit instruction. To test the hypothesis that CSs may more efficiently develop when they are explicitly recognized and practiced, Ex2 received direct instruction on CSs in addition to the activities used with Ex1. All groups received the same amount of instruction: five days a week, 90 minutes per day during the 11-week study. The time allotted for explicit CS instruction (Ex2) was approximately 10 minutes per day.

Teaching method for Ex1: Ss in Ex1 studied English using a series of tasks directing them to express meaning in English without anticipation of the linguistic forms they were supposed to use. Tasks were comprised of communicative exercises such as story telling, discussions, debates on different topics, and picture descriptions. (See Appendix for examples of the communicative activities.)

Teaching method for Ex2: In addition to using the same tasks as Ex1, Ss in Ex2 received explicit instruction about CSs. Analytic and holistic strategies are considered effective ways to convey meaning when the exact words to express the meaning are not immediately available (Poulisse, 1987). Therefore, these two types of CSs were selected for strategy training. On the first day of class, the Ss were introduced to definitions and examples of CSs. In addition, at the beginning of each class, Ex2 spent 5 to 7 minutes solving a lexical problem given on the blackboard. When the Ss encountered difficulty in expressing intended meaning, they were encouraged to solve the problems using the CSs they were studying.

Teaching method for the control group: The control group studied English in a form-focused learning context similar to traditional English classes in Japan. Instruction for this group focused on explicit explanation of particular linguistic forms followed by activities and tasks to practice those forms. The materials used with the two experimental groups (see Appendix) were adapted for use with the control group. With the audio-taped materials, different types of activities (multiple-choice, blank-filling, translation, dictation) were prepared. With the video materials, Ss were given the English transcripts and asked to translate them into Japanese. With the reading passages for debate, the major task for the control group was to translate the reading passages. Similarly, the pictures were used to teach formulaic expressions, with Ss tested on these formulaic expressions orally the following day.

Data Collection

Prior to (pre-test) and immediately after the instructional sessions (post-test), Ss were scheduled to individually perform two kinds of communicative tasks. Each session took approximately one hour. The first task was to describe ten concrete objects drawn on ten separate sheets of paper. Each object to be described was presented on a sheet of paper with a group of additional objects which shared characteristics with the target object and with one another. For example, a target object, a watering can, was presented with objects such as a bucket, a pitcher, a tin can, and a garbage can. Ss were told that native speakers of English would listen to their audio-taped descriptions to identify the objects described. Therefore, they were encouraged to describe the objects as specifically as possible, particularly if they did not know the exact words to name them. The second task was to narrate a story presented as uncaptioned cartoons. While the Ss were performing these tasks, the researcher was present and encouraged them to talk, by nodding or inserting words such as *Oh, I see, Say more*, and *Don't give up*. All performances on the two tasks were audio-taped and transcribed for subsequent analysis.

In addition, each subject's immediate retrospection on task performance was collected in Japanese, following the techniques proposed by Faerch and Kasper (1987). Immediately after the completion of a task, the Ss' audio-taped performances were played back and they were asked to describe what they were thinking while performing the task. The immediate retrospections were audio-taped, transcribed, and translated into English.

Based on the audio-taped and transcribed performances and the transcribed and translated immediate retrospection of the Ss, two raters categorized the Ss' use of CSs, according to the taxonomy described below.

Communicative Strategies Examined

The following CSs, based on existing typologies of CSs proposed by various researchers (Faerch & Kasper, 1983a; Poulisse, 1987; Tarone, 1983), were identified and scored in the analysis process of this study. The distinction between reduction strategies and achievement strategies was considered important in observing how closely learners' communicative goals were achieved (Varadi, 1983).

Taxonomy of Css

- Reduction Strategies
  - Message Abandonment
  - Meaning Replacement
- Achievement Strategies
  - Analytic
  - Holistic
  - Conceptual Transfer
  - Morphological Creativity
  - Linguistic Transfer

This study adopted Tarone's (1983) definition of message abandonment: "The learner begins to talk about a concept but is unable to continue and stops in mid-utterance" (p. 63). The following is an example of message abandonment: *When we hold the . . . I'm sorry I can't.* In this case, a subject intended to describe a *pot holder* but then gave up her attempt when she came across another unknown lexical item, *pot.*

Meaning replacement is different from message abandonment, as Faerch and Kasper (1983a) explain:

> [T]he learner, when confronted by a planning or retrieval problem, operates within the intended propositional content and preserves the topic but refers to it by means of a more general expression. (p. 44)

Based on this definition, meaning replacement strategies were identified when a subject did not entirely give up the problem-solving process, but the realized meaning was far from the communicative goal. For example, I *use this . . . use bake hot cake.* The subject attempted to convey the meaning *spatula*, by describing its functional characteristics, a tool to turn over hot cakes on a frying pan. In this process, the subject encountered other lexical problems such as *turn over.* Instead of giving up, the subject overgeneralized the meaning by simply using the word *bake.*

The category of achievement strategies was based on a typology proposed by Poulisse (1987). The first, analytic strategies, describes characteristic features of a referent to be expressed. An example of this strategy would be, *I use it to clean on the desk or bookshelf to get rid of the dust on the desk.* Not knowing the exact word *duster*, the subject tried to describe the object by presenting its functions. The second category, holistic strategies, are defined as tactics for manipulating a concept and referring to it by using the word for a related concept that shares similar features. For example, another student did not know the exact words *spiral shape*, so she substituted, *This is like a spring.* The next sub-category, conceptual transfer strategies, involve the application of an L1-based concept to refer to a concept in the target language. For example, while not knowing the word *sting* for a description of a wasp, one subject applied the Japanese concept of sting, *needle*, saying *this insect has needle and stick us.*

A fourth kind of achievement strategy (Poulisse, 1987) is morphological creativity, in which learners replace a morphological fragment with a creative one when they do not know the exact word representing a reference. For example, when unable to recall the word *pleased*, a student created a new word by adding a morpheme to the noun *pleasure*, saying *he is very pleasured.* Finally, linguistic transfer refers to the strategy of switching from the target language to the learner's first language to cope with a communicative problem.

In addition to the typology described above, CSs were also identified as operating either at the macro- or at the micro-level. When CSs were used to achieve a global communicative goal, they were considered to be operating at the macro-level; when CSs were used to compensate for a particular lexical item in the process of solving a global communicative problem, they were considered micro-level.

On Task 1, the global goal was to tell what the target objects were. CSs used to solve the problems at this level were regarded as being on a macro level, as in:

> *This is a tool to clean up a room. We drop a dust from our furniture from this tool. This tool have a long handle and the top that has a cloth or wing or something.*

In order to clarify the identity of the object *duster*, this subject described its function and appearance; applying macro-level CSs. On the other hand, to deal with the lexical problem *feather* in the description of the appearance of the duster, the subject used the word *wing*. This was regarded as a micro-level strategy.

On Task 2, CSs used to cope with difficulties in presenting a situation or an action were considered to be on a macro-level. In one of the cartoon stories, a man was lying in the space between two cars parked at the side of a street. The subject, when unable to express the meaning *a man was lying down on a space*, omitted this meaning unit. This strategy was considered macro-level. On the other hand, some CSs used to compensate for particular lexical items such as the use of the general term *doctor*, in place of *archeologist* were considered micro-level.

All CSs used by a subject to realize an intended meaning were counted separately. Two raters independently identified and classified CSs on the basis of their common characteristics, following the taxonomy of CSs established for this study. The results of the raters' identification and scoring showed reasonably good agreement. An average of 73 percent inter-rater reliability was obtained. For those instances in which agreement was not reached, a face-to-face meeting of the raters was held to resolve the difference.

Evaluation of Subjects' Communicative Performance

Besides the identification of CSs, Ss' audio-taped communicative performances were evaluated by the two raters. On Task 1, raters were asked to identify the items described by the Ss. The effectiveness of each subjects' performance was evaluated according to the number of correct objects identified by the raters.

On Task 2, the raters were asked to evaluate the subject's audiotaped communicative performances holistically, on a scale ranging from 0 to 6, based on the amount of relevant information provided by the subject. Inter-rater reliabilities on this evaluation averaged 80 percent, which was considered reasonably high.

Results

The discussion of results focuses on discussion of message abandonment, meaning replacement, analytic, and holistic strategies. Among the CSs listed in the taxonomy for this study, a number of strategies—conceptual transfer, morphological creativity, and linguistic transfer—were infrequently used, both on the pre-test and the post-test. Their sporadic use suggested a preference of an individual subject rather than that of a group. Therefore, use of these strategies will not be discussed here.

Tables 1 – 3 display the raw frequency counts of those strategies used by the Ss, along with means and standard deviations for both Task 1 and Task 2.

Table 1: Descriptive Statistics of Combined Reduction Strategies and Combined Analytic and Holistic Strategies for Task 1 and Task 2

	Task 1				Task 2			
	MA + MR (macro)		A + H (micro)		MA + MR (macro)		A + H (micro)	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Ex1 Frequency	30	7	20	73	29	6	1	24
M	6	1.4	4	14.6	5.8	1.2	0.2	4.8
SD	1	1.14	1	1.6	1.3	0.8	0.4	2.0
$n = 5$								
Ex2 Frequency	34	3	20	81	31	0	3	48
M	6.8	0.6	4	16.2	6.2	0	0.6	9.6
SD	2.59	0.89	2.65	3.42	1.4	0	0.5	2.5
$n = 5$								
CG Frequency	27	31	20	27	32	38	3	13
M	5.4	6.2	4	5.4	6.4	7.6	0.6	2.6
SD	2.3	0.84	2.55	1.95	1.5	3.28	0.8	2.07
$n = 5$								

MA = message abandonment, MR = meaning replacement, A + H = analytic and holistic strategies

Table 2: Descriptive Statistics of
Individual Reduction StrAtegies for Task 1 and Task 2

	Task 1				Task 2			
	MA + MR (macro)		A + H (micro)		MA + MR (macro)		A + H (micro)	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Ex1 Frequency	11	0	19	7	23	1	6	5
M	2.2	0	3.8	1.4	4.6	0.2	1.2	1
SD	0.84	0	0.83	1.14	0.54	0.44	0.83	0.7
$n = 5$								
Ex2 Frequency	18	0	16	3	28	0	3	0
M	3.6	0	3.2	0.6	5.6	0	0.6	0
SD	3.51	0	2.48	0.89	1.34	0	0.54	0
$n = 5$								
CG Frequency	15	11	12	20	24	25	8	13
M	3	2.2	2.4	4	4.8	5	1.6	2.6
SD	1.87	1.3	1.67	1.58	1.30	3.74	1.14	1.51
$n = 5$								

MA = message abandonment, MR = meaning replacement, A + H = analytic and holistic strategies

Table 3: Descriptive Statistics of Individual Analytic and Holistic Strategies

	Task 1				Task 2			
	MA + MR (macro)		A + H (micro)		MA + MR (macro)		A + H (micro)	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Ex1 Frequency	17	62	3	11	0	7	1	17
M	3.4	12.4	0.6	2.2	0	1.4	0.2	3.4
SD	1.51	1.51	0.89	1.30	0	0.89	0.44	1.5
$n = 5$								
Ex2 Frequency	16	65	4	15	1	12	2	36
M	3.2	13	0.8	3	0.2	2.4	0.4	7.2
SD	2.28	3.80	0.44	1	0.44	0.89	0.54	1.9
$n = 5$								
CG Frequency	15	20	5	7	0	1	3	12
M	3	4	1	1.4	0	0.2	0.6	2.4
SD	2.35	1.58	0.71	0.89	0	0.44	0.89	2.07
$n = 5$								

MA = message abandonment, MR = meaning replacement, A + H = analytic and holistic strategies

BEST COPY AVAILABLE

Task 1

Statistical results: All three groups' pre-test scores of the combined macro-level message abandonment and meaning replacement strategies, which were used for problems of target concepts, show no statistically significant difference ($F=0.569$; $df=2, 12$; $p>0.5$). On the post-test, in contrast, the two experimental groups showed a substantial drop in the use of these strategies, while the control group stayed within the range of the pre-test scores. A one-way ANOVA shows that these differences are statistically significant ($F=49.14$; $df=2, 12$; $p<0.01$). Furthermore, the additional comparison between the two experimental groups showed no statistically significant difference ($t=1.24$; $df=8$; $p>0.05$). Individual macro-level reduction strategies show the same trend as the combined macro-level reduction strategies (see Table 2).

The data for macro-level analytic strategies complemented those for the reduction strategies. The means ranged from 3.0 to 3.4 on the pre-test, while on the post-test they ranged from 4.0, for the control group, to 12.4 and 13.0 for Ex1 and Ex2 (see Table 3). Although the control group increased slightly from pre- to post-test, there was a three-fold increase for both Ex1 and Ex2. A one-way ANOVA on the post-test shows this difference to be statistically significant ($F=19.67$; $df=2, 12$; $p<0.01$).

Descriptive results: Reduction strategies were the primary strategies used across the three groups on the pre-test on Task 1. Nevertheless, this does not mean that the Ss did not have knowledge of effective CSs. Their immediate retrospection reveals that they went through the problem-solving process by analyzing the characteristics of the objects. However, they often stopped in the middle of these processes, probably because the communicative problems they experienced seldom appeared in isolation but tended to be interlocked with one another. One subject's performance on Task 1 on the pre-test illustrates. *When we hold the . . . I'm sorry I can't.* The subject's retrospection reveals that she intended to describe the function of the target object, *pot holder*. However, as soon as she came upon another problem with a lexical word *pot*, she gave up her attempt.

On the post-test of Task 1, similar message abandonment behavior was observed among the Ss of the control group, as in this example of a subject's attempt at describing a cylinder: *bottom is circle . . .* (silence). In contrast, Ex1 and Ex2 Ss drastically increased the use of analytic or holistic strategies on the post-test to describe the particular object more specifically and accurately, as seen in the following examples:

> *(Ex1) The shape is . . . like drink can . . . can drink . . . the bottom . . . object . . . the bottom is circle . . . when you see the side . . . the shape is . . . rectangle.*
>
> *(Ex2) The bottom . . . the bottom and top is circle . . . circle shape . . . and can juice is . . . the same of this shape.*

Task 2

Statistical results: The results concerning macro-level message abandonment strategies, which were used with problems of propositional meaning such as an action and a situation, show similar patterns to those obtained on Task 1. Message abandonment strategies, which constitute the predominant strategies on the pre-test across the three groups, shrank significantly on the experimental groups' post-test performance, while the control group remained within the range of the pre-test scores ($F=8.46$; $df=2, 12$; $p<0.01$). In addition, a t-test showed no statistically significant difference between Ex1 and Ex2 ($t=1.00$; $df=8$; $p>0.05$). On the other hand, there were statistically significant differences among the three groups on the macro-level meaning replacement strategies used on the post-test. A one way ANOVA shows that the Ex2, which had strategy training, eliminated the use of this strategy on the post-test compared to the Ex1, which did not have strategy training, as well as the control group ($F=9.24$; $df=2, 12$; $p<0.01$; $t=3.16$; $df=8$; $p<0.05$).

The results of micro-level achievement strategies, which were used to solve lexical problems within a process of solving a problem of propositional meaning such as an action and a situation, reveal a striking increase in their use among the experimental groups on the post-test. Ex2, which received strategy training, used both micro-level analytic and holistic strategies more than three times as often as the control group and twice as often as Ex1, though no difference was observed among the three groups on the pre-test. A one-way ANOVA as well as a t-test showed the differences among all three groups to be statistically significant ($F=12.97$; $df=2, 12$; $p<0.01$; $t=3.31$; $df=8$; $p<0.05$). The results of individual strategies show similar patterns, though raw frequency counts of holistic strategies were found to be higher than analytic strategies on Task 2 (Analytic: $F=9.33$; $df=2, 12$; $p<v0.01$; Holistic: $F=9.34$; $df=2, 12$; $p<0.01$).

Descriptive results: Reduction strategies, especially message abandonment, were the primary strategies applied by all three groups on the pre-test. While the control group's use of these strategies remained at the same level as in the pre-test, the experimental groups demonstrated strategic change in their communicative performance on the post-test, with significant increase of their use of micro-level analytic and holistic strategies.

The Ss' performance in the story narration task indicates that the success rate of solving lexical problems depends on top-down as well as bottom-up processing. When solving communicative problems, Ss in the meaning-focused learning context tended to analyze meaning units within the context of the particular communicative task in order to choose an appropriate CS. This context-dependent approach enabled Ss of the experimental groups to express their meaning more accurately and effectively than Ss of the control group, who relied primarily on a context-free approach. One part of the cartoon story on Task 2 was as follows.

> An archaeologist discovered an ancient document on which a statue was drawn. Assuming that it was academically valuable, he decided to search for the statue. After he and his followers endured hardship, they managed to reach their destination, where there were ruins. The archaeologist climbed up one monument and found the statue for which he had been searching.

At this point in the story, there appeared several objects (e.g., *statue, ruins, monument*) which shared some characteristics in this context and therefore needed to be described distinctively in order to make the story coherent. However, most Ss in the control group failed to do so. Following is an example of a control group subject's performance:

> *A man who studied . . . who study . . . old monument . . . he found . . . monument . . . in . . . some . . . place. So, he . . . go to find it . . . he gathered a lot of people to . . . find it with him . . . at last they find a monument . . . and he find the monument he studied.*

In this part of the subject's narration, she used the word *monument* to refer to the statue in the story. However, since there were both statues and monuments in the story, her use of the word was confusing. Rather than just a lack of vocabulary, her problem seems to be also a lack of analysis of the various meaning units within the global context. Objects such as an ancient painting, a figure drawn on it, ruins, monument and statue all were present in the story and therefore should have been referred to with distinct, specific vocabulary items. In fact, Ss of the control group generally failed to analyze and reconstruct meaning elements to represent particular objects in the context. As a result, their performance produced inconsistency and confusion in the story-telling task.

On the other hand, Ss in the meaning-focused learning context managed to distinguish the objects by using different vocabulary items and thus maintained coherence and clarity in their narration of the story. For example, *a figure* was described as *a doll* or *a monkey. Monument* was described as *pyramid, castle,* or *building made of stone*. The following are two segments of an experimental group subject's narration:

> (Ex1) *One famous doctor . . . find . . . a paper. This paper has drawn . . . the doll. He think . . . this is a god . . . of . . . Inca people—ancient Inca people believed . . . God. So, he think . . . in Inka, there is a doll like this . . .*
>
> (Ex2) *There is one man and he studied ancient matter very well and . . . a certain time . . . he found . . . very old picture and . . . there . . . the picture like money . . . is drawn, so he thought . . . South America . . . Ahh . . . there is . . . this picture- . . . he thought . . . same . . . same object . . . he thought there must be same object in South . . . America . . .*

These results suggest the importance of an interactive operation between higher-order interpretive skills at the discourse level and lower-order lexical knowledge in the process of solving lexical problems.

Evaluation of Subjects' Communicative Performance

The independent ratings of identified objects on Task 1 by two raters show that both raters, listening to the Ss' audio-taped description of the target objects without knowing to which group they belonged, could identify two to three out of the 10 objects on the pre-test. In contrast, on the post-test, the same raters could identify an average of eight objects for Ex1 and nine objects for Ex2. These results contrast with the number of objects identified by the same raters from the control group's description on the post-test where the average number remained three objects.

Similar results were obtained from the Ss' communicative performance on Task 2. On this task, the Ss' audio-taped performances were holistically evaluated by the same raters on a scale that ranged from 0 to 6, according to the amount of relevant information provided by the Ss. The average scores given by the raters on the pre-test were 2.6, 2.4, and 2.4 for Ex1, Ex2, and the control group, respectively. Though the control group's scores on the post-test stayed within the pre-test range (2.5), those on the experimental groups' performance improved significantly. The raters scored an average 4.3 for Ex1 and 4.5 for Ex2. One-way ANOVAs yield statistically significant differences between the control group and the experimental groups (rater 1: $F = 26.74$; $df = 2, 12$; $p < 0.01$; rater 2: $F = 15.96$; $df = 2, 12$; $p < 0.01$). On the other hand, no statistical significance was obtained in the comparison between Ex1 and Ex2.

The results of the evaluation of the Ss' communicative performance positively correlated with their use of CSs. The experimental groups' macro-level analytic strategies on task 1 correlated with the success rates in which the objects they described were identified by the two raters. Similarly, the experimental groups' use of micro-level analytic and holistic strategies resulted in perceptible improvement in the evaluation of effectiveness of their communicative performance on Task 2 on the post-test.

Discussion and Conclusions

The purpose of this study was to examine whether different types of learning contexts would contribute to variation in the use of CSs by learners on communicative tasks. Also, in order to examine whether direct strategy training in a meaning-focused learning context is necessary for learners to develop effective CSs, explicit strategy training was provided for Ex2.

Post-test results did not show generally significant effects of this training, except on the Ss' uses of macro-level meaning replacement and micro-level holistic strategies on Task 2. Ss of Ex2 did pay more attention to the meaning to express, but according to the raters' evaluation there was no significant difference between the performance of Ex2 and that of Ex1 on Task 1 and 2. Therefore, the overall effect of direct strategy training upon learners' choice of CSs was found to be modest. The learning context, then, seems to be responsible for the major effect on the experimental Ss' strategy choice and application.

As a whole, reduction strategies were the primary strategies used by the three groups on the pre-test. In contrast, on Task 1 on the post-test, macro-level analytic strategies became the major strategies for Ex1 and Ex2. On Task 2, micro-level holistic strategies were the most commonly used type of strategy by these two groups. The significant increases in the use of these strategies seem to correlate with the increased effectiveness of these groups' communicative performance. On the post-test, more objects described by Ss of the experimental groups were identified by the raters, and the communicative performance of both groups was judged to be more effective than that of the control group.

It is doubtful that these results were due to the development of the experimental groups' grammatical or lexical knowledge of English. The two groups' task performance (a few examples have been provided previously) clearly indicates that these students had no more linguistic knowledge than the Ss in the control group.

These findings provide further evidence, as suggested by cognitive views of L2 acquisition (Bialystok, 1990; Bialystok & Sharwood-Smith, 1985; Ellis, 1986; O'Malley & Chamot, 1990; Tarone, 1983), that L2 learners' language ability does not consist of linguistic knowledge alone, but develops as a result of interaction between linguistic knowledge and cognitive processes. Furthermore, the process of applying a particular CS does not operate only at the local lexical level but also includes the analysis of a global communicative goal and the establishment of meaning units within the goal. These findings suggest that strategies used to cope with communicative problems are not automatically transferred from the learner's first language but are acquired in the process of using the target language in particular contexts. While learners may possess effective CSs in their first language, this does not guarantee their being able to apply these strategies to problem-solving in target-language communication.

These findings provide some implication for language pedagogy. If the goal of a language program is to develop learners' communicative ability, a learning context which focuses on explicit instruction of linguistic structures alone may not be sufficient for reaching such a goal. In such a context, learners are likely to develop strategies such as analyzing linguistic structures and memorizing bits of linguistic information but may fail to develop strategies of retrieving those linguistic resources from memory by analyzing a global communicative goal and constructing meaning units within the goal. This may be especially true when the target language is being learned in a foreign language environment where the classroom is the primary source of input for the learners. Therefore, to help learners develop their communicative ability of the target language, various instructional procedures need to be considered.

First, the curriculum should be constructed with a clear goal, as seen in task-based language instruction (Nunan, 1988, 1989), allocating sufficient instructional time to tasks directing learners to engage in problem-solving processes to convey their intended meaning. Second, teaching materials should be chosen and developed with a sound theoretical basis, corresponding with the instructional goal. Instructional materials should include visual aids such as pictures, maps, and symbols which are useful for creating communicative tasks. Video materials also initiate communicative language use such as discussions and debates. Finally, the teaching methods to be adopted must go hand in hand with the goals set up by the curriculum. Learners as active participants in the learning process should be placed in the center of learning, allocating sufficient time for letting them engage in communicative tasks in paired work or small group work.

Finally, in terms of future direction for studies on CSs, the necessity and effect of strategy instruction need further investigation. Some researchers (Bialystok, 1990; Bongaerts, Kellerman & Bentlage, 1987;

Bongaerts & Poulisse, 1989) question the necessity of teaching CSs to learners, from a standpoint that L2 learners already possess those CSs in their native languages and therefore are able to transfer them to L2 communication. On the other hand, some studies (Dornyei, 1995; Wildner-Bassett, 1986) reveal a positive effect of strategy training on learners' communicative performance.

This study, with the inclusion of a strategy training group, examined whether or not a learning context alone is a necessary and sufficient condition for learners to develop effective CSs. The obtained results show no difference between a group with strategy training and the one without it, except on the use of micro-level analytic and holistic strategies on the narration task. However, this was a preliminary study with small sample size and with a short period of time for strategy training. More studies with larger sample size and with more extended period of strategy training need to be conducted before reaching any conclusion. Furthermore, such studies need to incorporate variables influencing the outcome such as learners' proficiency level, learners' personality traits, qualitative aspects of fluency, and different types of discourse in which learners engage.

Ryu Kitajima, Ph.D. Second Language Education, teaches Japanese as a foreign language at San Diego State University.

Note

1. This paper is based on part of the author's doctoral dissertation, which was accepted by SUNY at Buffalo in 1993. An earlier version of this paper was presented at the 28th TESOL convention in Baltimore, March 8-12, 1994.

References

Bialystok, E., & Sharwood-Smith, M. (1985). Interlanguage is not a state of mind: An evaluation of the construct for second language acquisition. *Applied Linguistics, 6*, 101-117.

Bialystok, E. (1990). *Communication strategies*. Cambridge: Basil Blackwell.

Bongaerts, T., Kellerman, E., & Bentlage, A. (1987). Perspective and proficiency in L2 referential communication. *Studies in Second Language Acquisition, 9*, 171-200.

Bongaerts, T., & Poulisse, N. (1989). Communication strategies in L1 and L2: Same or different? *Applied Linguistics, 10*, 253-268.

Chen, S. (1990). A study of communication strategies in interlanguage production by Chinese EFL learners. *Language Learning, 40*(2), 155-187.

Corder, S. (1983). Strategies of communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 15-19). London: Longman.

Dornyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly, 29*, 55-85.

Ellis, R. 1986. *Understanding second language acquisition*. Oxford: Oxford University Press.

Faerch, C., & Kasper, G. (1983a). Plans and strategies in foreign language communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 20-60). London: Longman.

Faerch, C., & Kasper, G. (1983b). On identifying communication strategies in interlanguage production. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 210-238). London: Longman.

Faerch, C., & Kasper, G. (1987). From product to process: Introspective methods in second language research. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 5-23). Clevedon: Multilingual Matters.

Nunan, D. (1988). *Syllabus design*. Oxford: Oxford University Press.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

O'Malley, M., & Chamot, A. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Paribakht, T. (1985). Strategic competence and language proficiency. *Applied Linguistics, 6*, 132-146.

Poulisse, N. (1987). Problems and solutions in the classification of compensatory strategies. *Second Language Research, 3*, 141-153.

Poulisse, N., & Schils, E. (1989). The influences of task- and proficiency-related factors on the use of compensatory strategies: a quantitative analysis. *Language Learning, 39*, 15-48.

Tarone, E. (1983). Strategies of target language learner communication: Message adjustment. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 61-74). London: Longman.

Varadi, T. (1983). Strategies of target language learner communication: Message adjustment. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 79-99). London: Longman.

Wildner-Bassett, M. (1986). Teaching and learning 'polite noises': Improving pragmatic aspects of advanced adult learners' interlanguage. In G. Kasper (Ed.), *Learning, teaching and communication in the foreign language classroom* (pp. 163-178). Aarhus: Aarhus University Press.

(Received April 10, 1996; revised July 25, 1996)

The *Eiken* Test: An Investigation

Laura MacGregor
Sapporo International University

The *Eiken* tests, first administered in 1963 by the Society for Testing English Proficiency (STEP), are highly respected in social, educational, and employment circles and taken by millions each year. However, upon closer scrutiny, it appears that the *Eigo Kentei Kyokai* (*Eikyo*) operates on its own terms. Unlike TOEFL and TOEIC, *Eikyo* does not make information about the tests' reliability or validity available to the public. Therefore, important questions remain unanswered: Are the *Eiken* tests reliable and valid instrument? Do the *Eiken* tests really function as tests of English proficiency? This paper examines the *Eiken* pre-second level test from June 1994. The test was administered to 168 first-year Japanese college students. The results provided data for reliability and validity studies, in an effort to shed light on the value of the *Eiken* pre-second level test as a reputable test instrument of English proficiency. The results of the studies conducted here are far from encouraging with regards to the *Eiken* pre-second level test's reliability and validity.

英検は1963年から英語検定協会によって実施されており、社会的、教育的に、また企業によっても高く評価され、毎年数千人の受験者がいる。しかしながら詳しく調査してみると、英検は独自の方式で実施されており、TOEFLやTOEICと違って信頼性や妥当性についての情報が公開されていない。従って、以下のような重要な疑問は答えのないままである。英検は信頼性と妥当性のあるテストであるのか。英検は本当に英語能力のテストとして機能しているのか。本研究は、英語能力を測定するテストとしての英検の価値を知るために、1994年6月に実施された英検2級のテストを168人の日本人大学一年生に実施し、その信頼性と妥当性を調査した。その結果、英検の信頼性と妥当性ははなはだ心細いものであることがわかった。

Japan is a country whose people thrive on tests, from *kendo* to calligraphy, flower arranging to gift wrapping; tests which evaluate almost every skill imaginable are available for the taking, so to speak. Walk into any bookstore or culture center and you'll see an array of posters and pamphlets advertising such tests. In the academic world, tests are in abundance as well. Entrance examinations which determine students' future high school and post-secondary careers are a fact of life for virtually every family.

JALT Journal, Vol. 19, No. 1, May, 1997

By far, the oldest and best established English language tests in Japan are the *Eiken* tests (*Eigo Kentei*), produced by *Nihon Eigo Kentei Kyokai* (*Eikyo*), or in English, STEP: the Society for Testing English Proficiency. Since it began offering the *Eiken* tests in 1963, *Eikyo* has enjoyed a long period of unprecedented success: The Ministry of Education endorses the *Eiken* tests,[1] and recommends students take them. Some schools even offer courses dedicated to *Eiken* test preparation. In the working world, it has made its mark as well. Many employers regard *Eiken* test qualification as a valuable asset and look for it on prospective employees' resumes (MacGregor, 1995). However, despite the test's wide acceptance and use, one important question has been overlooked: Are the Eiken tests reliable and valid instruments to measure English proficiency? Eikyo has failed to give a direct answer. Other language tests, such as TOEFL and TOEIC, publish regular reports with statistical analyses of reliability and validity (TOEFL, 1995; TOEIC, 1995; Woodford, 1980, 1992). Why doesn't *Eikyo*? This question has been raised by other language educators concerned about Eikyo's position as well (Bostwick, 1995; Brown, 1995; Gorsuch, 1995).

This paper seeks answers to the above questions of reliability and validity by conducting analyses on a set of data collected by the author. Because certain reliability analyses are best suited to certain types of tests, it is necessary to begin with the question, is the *Eiken* a criterion-referenced test (CRT) or a norm-referenced test (NRT)? To help answer this question, various aspects of the *Eiken* test will be compared with those of TOEFL and TOEIC. This discussion is followed by an investigation of the *Eiken* pre-second level's[2] reliability and validity, which focuses on four questions: 1) Is the *Eiken* test appropriate for the group *Eikyo* claims to evaluate? 2) Do the test items reflect the practical English found in daily life that *Eikyo* claims to test? 3) Does the test measure the abilities that *Eikyo* claims it does? and 4) Are there any poorly constructed test items? An examination of how the *Eiken* test is scored and how the scores are reported follows. Finally, recommendations are made as to how the *Eiken* test can better serve students and teachers.

Background

The Society for Testing English Proficiency (STEP) was established over 30 years ago as part of a plan by Japan's Ministry of Education to develop education across the nation. Specifically, the goals of the STEP were to popularize and improve the level of practical English in Japan (*Nihon Eigo Kentei Kyokai*, 1994a). In 1963, the same year that the TOEFL was first administered, the *Eiken* tests were given for the first time to about 37,000 people at three levels (first, second, and third). Five years later, the *Eiken* tests received official approval from the Ministry of Education, which began actively promoting the *Eiken* tests as important tests of English proficiency. With this stamp of approval, the number of test-takers soared, and continues to increase each year.[3] The original three-level test has grown to seven levels: first, pre-first, second, pre-second, third, fourth, and fifth, the first level being the most difficult. The most recent addition to this series was the pre-second level in 1994, introduced to bridge what was felt to be a wide gap in difficulty between the third and second levels.

For levels one to three, the tests are given in two stages. The first stage is a written test (reading and listening comprehension) and the second stage is a speaking (interview) test. Both stages are offered twice a year, in June and October.[4] The focus of this paper is on the first stage of the pre-second level *Eiken* test of June, 1994 (Nihon Eigo Kentei Kyokai, 1994b).

What kind of tests are the Eiken tests?

In the literature on language testing, two types of tests are found (Brown & Yamashita, 1995; Henning, 1987; Hughes, 1989), criterion-referenced tests (CRTs) and norm-referenced tests (NRTs). The differences between the two types lie, not in the actual items themselves, but in the purpose of the tests, how the tests are scored, and how the test scores are used. Therefore, just looking at the test instrument is not enough to determine what kind of test it is.

The purpose of a criterion-referenced test is to evaluate how well the test-taker can perform a specific set of tasks. For example, classroom and term tests evaluate how well a student has learned a defined set of material over a specific period of time. Brown (1996) explained that, "the interpretation of scores on a CRT is considered absolute in the sense that each student's score is meaningful without reference to other students' scores" (p. 2). Therefore, CRT scores do not necessarily conform to a normal distribution.

A norm-referenced test, on the other hand, measures general language abilities. Each student's score is interpreted relative to the scores of all the other students who took the test and the scores generally fall along a normal distribution curve. The TOEFL and TOEIC are norm-referenced tests. The purpose of the TOEFL is to evaluate the English proficiency of foreign nonnative speakers of English, primarily those who intend to study at colleges and universities in the United States or

Canada. Thus, the content of the TOEFL focuses on English for academic purposes. The TOEIC (Test of English for International Communication) is an English language proficiency test which measures how well non-native speakers of English can communicate in English with others in business, commerce, and industry.

Like TOEFL and TOEIC, the *Eiken* tests are English proficiency tests for non-native speakers of English. However, the *Eiken* tests are different in at least two ways. First, the *Eiken* tests are not just one test, but seven different tests. These divisions allow the test-makers to clearly define the material covered, a characteristic of a CRT. *Eikyo* described the contents of its pre-second level test, as follows:

> Successful examinees are able to understand and use general English needed in daily conversation. (High school level; appropriate for a wide range of ages, from high school students to adults in Japan.)
>
> The successful examinee is:
>
> (1) Able to converse about common daily topics. (Able to conduct simple business by telephone; to make easy explanations, leave messages, do simple interpretation, etc.)
>
> (2) Able to read material about common everyday topics. (Able to read news articles, letters, simple pamphlets, etc.)
>
> (3) Able to write about common everyday topics. (Able to write simple letters, notes, memos, etc.) (*Nihon Eigo Kentei Kyokai*, 1994a, p. 8)

Dividing the tests into seven levels is a practical way of handling the wide population, from junior high school to post-graduate levels, that *Eikyo* tests. Further, if students take and pass a test at a level appropriate to their ability, that success is seen to be motivational for their continued study.

The second point in which *Eiken* tests differ from TOEFL and TOEIC is its method of reporting scores. *Eiken* uses a pass/fail reporting system while TOEFL and TOEIC use a converted scale-score reporting system, which makes these two tests "'user-defined' in that scores can be considered in a variety of ways depending upon the requirements of a particular individual or client" (Wilson, 1993, p. 2). Although the reporting styles differ, all these tests follow NRT procedures by using some form of statistical analysis to translate raw scores into standard scores. In other words, none of the tests report their scores as absolute scores.

The question, "What kind of tests are the *Eiken* tests?" remains unanswered. According to Bostwick (1995), "the *Eiken* STEP claims to be a criterion-referenced test in that it specifies proficiency standards and attempts to identify whether the student can pass the pre-established standard" (p. 58). The fact that the *Eiken* tests are divided into seven levels, the purpose of each level clearly defined by a set of specific tasks, with the language skills required to pass each level specifically defined, gives it qualities of a CRT. However, the way the test is scored (i.e., by converting raw scores to standard scores) is characteristic of an NRT. Therefore, it makes sense to call the *Eiken* tests hybrid CRT/NRT. Knowing that the tests are scored as NRTs is helpful in choosing appropriate reliability measures.

The Study

Method

Materials: The pre-second level test was originally developed for 2nd and 3rd-year high school students (16 and 17-year-olds) (*Nihon Eigo Kentei Kyokai*, 1994c). The most recent statistics show that this group forms the majority of test-takers. For the June, 1996 test, 75% of those who took the pre-second level (227,666 out of a total of 303,955) were senior high school students, 4% (12,471) were junior college students, and 3% (8,549) were university students (*Nihon Eigo Kentei Kyokai*, September, 1996).

There are 75 multiple-choice items on the pre-second level written test. Part 1, which tests vocabulary, idioms, grammar, usage, and reading composition, has 55 items, and Part 2, the listening section, has 20 items. Each item is worth one point for a total of 75 points.

Subjects: The subjects for this study were 182 first-year students (ages 18-20) in five classes at a junior college. Although the reports by *Eikyo* indicate that the pre-second level test is ideally suited to high school students, this higher age level was selected as it best matched the general ability and experience of the group, as outlined below.

A survey, in Japanese, accompanying the *Eiken* pre-second level test to determine the students' experience in taking it, showed that 17% of the students had tried the pre-second level test at least once before but failed, while 40% had never taken an *Eiken* test before. However, because the format of the *Eiken* pre-second level test is similar to high school English tests, it was concluded that lack of *Eiken* experience would not adversely affect the data. Fully 43% had previously passed the third level, confirming that the pre-second level was the most appropriate for this group.

Procedure: In May, 1996, the pre-second level test of June, 1994 was administered to all 182 Ss, along with the survey of *Eiken* experience. After a review of the results of the survey, the results for 14 Ss who had previously passed pre-second level were eliminated from consideration. The remaining papers (N=168) were scored by hand. The test had 75 items, each worth one point.

Analysis

Analyses to evaluate test reliability were done on the data collected in three categories: Descriptive statistics, item statistics, and consistency estimates. Analyses to evaluate validity were done by comparing the contents of the test with the aims set out by *Eikyo* and the course of study for high school English education. The construction of the individual test items was also examined.

Reliability

Descriptive Statistics: Minimum score, maximum score, midpoint, mean, and standard deviation were calculated. The midpoint is the score which is halfway between the highest and the lowest score. The midpoint, together with the mean, are two statistics which help locate the middle or typical score (Brown, 1996, p. 109).

Item Statistics: Investigation into the reliability of the *Eiken* test began with a look at two types of item statistics: item facility and item discrimination. Item facility (IF) is "a statistical index used to examine the precentage of students who correctly answer a given item" (Brown, 1996, p. 64). The following formula for item facility was used to evaluate individual test items:

$$\text{IF} = \frac{N_1 \text{ (number of examinees who answered correctly)}}{N_2 \text{ (number of examinees who took the test)}}$$

Item discrimination (ID) "indicates the degree to which an item separates the students who performed well from those who performed badly" (Brown, 1996, p. 66). In order to calculate the ID index, it is necessary to differentiate the high scorers from the low scorers. In this study, the upper and lower thirds (33%, or 56 students each) were taken to represent the high scorers and the low scorers respectively. The ID was calculated as follows:

$$\text{ID} = \text{IF (item facility of high scorers)} - \text{IF (item facility of low scorers)}$$

That IF and ID are closely connected is apparent from the above formula: If the item is easy (i.e. has a high IF), there should be little discrimination (low ID), and if the item is rather difficult, the discrimination should be high. If the item is too difficult, there should be no discrimination. According to Brown (1996, p. 69), ideal items in an NRT project have an average IF of .50 and the highest available ID. However, in reality, "items rarely have an IF of exactly .50, so that those that fall in a range between .30 and .70 are usually considered acceptable" (Brown, 1996, pp. 69-70). The items outside this range should either be set aside for revision or discarded. Examination of the ID of the remaining items further evaluates their suitability. The following guidelines were used to evaluate item discrimination:

.40 and up -	very good item
.30 - .39 -	reasonably good, but possibly subject to improvement
.20 - .29 -	marginal item, usually needing and being subject to improvement
below .19 -	poor item, to be rejected or improved by revision (Ebel, 1979, p. 267, cited in Brown, 1996, p. 70)

Consistency Estimates

In this study, the Kuder-Richardson Formula 20 (KR-20) reliability estimate and the standard error of measurement (SEM) were used (Brown, 1996, p. 199, p. 207) to estimate the *Eiken* pre-second level test's reliability. KR-20 was chosen over two other methods of calculating reliability, KR-21 and Cronbach's alpha, for two reasons: 1) it is reported to be the most accurate of the three (Brown, 1996, p. 199), and 2) the results could be compared with the reliability statistics of the TOEFL and TOEIC, which also use KR-20. Further, the *Eiken* pre-second level test follows criteria required by KR-20, in that each item is worth one point and is scored as correct/incorrect.

Test reliability is important because it measures the consistency of the test instrument. If a student takes a test on one day, and then takes the same test again a week or two later, it should produce nearly identical results, that is, if the test is a reliable instrument and little or no learning has taken place between the first and second testing. A test with a reliability coefficient of 1.0 would give precisely the same results for a particular group of test-takers regardless of when it was administered: it would be 100% reliable. Therefore, it is the goal of test-makers to attain the highest possible reliability coefficient.

The standard error of measurement (SEM) is another useful statistic for estimating reliability of NRTs. According to Brown (1996, p. 206), the SEM is "used to determine a band around a student's score within which that student's score would probably fall if the test were administered to him or her repeatedly." A multitude of factors affect test performance, only a few of which are connected with the test instrument itself, including items not matched to the purpose of the test, formats unfamiliar to the test-taker, and poorly constructed test items. The majority of factors affecting test performance are related to physical setting (i.e. the test room), and the mental and physical condition of the test taker. With all of these potential distracters, it is sensible to factor in the SEM when determining the cutoff scores.

Validity

A test is a valid instrument if it measures accurately what it claims to measure. For example, an arithmetic test of addition should contain only test items which ask students to add numbers. Further, the test items should not contain ambiguities or misleading information. In the case of the addition test, we can tell whether it is a valid instrument or not just by looking at it. Assessing the validity of English proficiency tests like the *Eiken* is a bit more complicated.

Validity analyses are done on a regular basis by TOEIC and TOEFL. According to a TOEIC report:

> The first validity studies involved the administration of TOEIC to a representative population of Japanese managers, technicians, bankers, and other employees who require English in their work. Researchers compared the candidates' performance on TOEIC to their performance on direct measures of listening, speaking, reading, and writing and determined the correlations. (TOEIC, 1995, p. 3)

Therefore, TOEIC compares its test results to direct four-skills test results to determine test validity. TOEFL conducts a similar type of validity analysis: "TOEFL validation is based upon correlations between test performance of foreign students studying in U.S. colleges and universities and their performance in degree-granting educational programs" (TOEIC, 1995, p. 3). Both TOEFL and TOEIC test for what Woodford calls "concurrent" validity:

> If a language test is supposed to measure whether a person can read Japanese or not then the person who scores high on the test should be able to pick up the Japanese newspaper and tell us what the lead article says. The low scorer should not be able to do it. (1980, p. 4)

It is not known whether *Eikyo* is also doing such validity studies.

Other validation studies involve comparing the results of one test with the results of another. In 1993, TOEIC (Wilson, 1993) published a report of a research study which linked the TOEIC listening section scores to the scores of the Language Proficiency Interview, a direct assessment of oral language proficiency developed by the Foreign Service Institute of the U.S. Department of State. The numerical correlations between LPI and TOEIC listening sections (.83) proved to be consistently high, suggesting that both tests are, as they claim, effective measures of the ability to understand and use spoken English.

Content validity compares the test specifications with the test contents. If the individual test items match the specifications, then the test can be said to have content validity. This is a subjective evaluation which should be done by a group of testing experts. For the present study, the resources necessary to do the type of validity research described above were not available. Instead, four general questions pertaining to test validity were posed: 1) Is the pre-second level test really appropriate for the group *Eikyo* aims to examine? 2) Do the contents of the test items reflect aspects of "daily life" in Japan, as *Eikyo* claims (*Nihon Eigo Kentei Kyokai*, 1994a)? 3) Do the test items really measure the abilities that they purport to? and 4) Are there any poorly constructed test items?

Results and Discussion

Reliability

Descriptive Statistics: Table 1 shows the descriptive statistics. The midpoint and mean indicate that the typical scores for those who took the test were just above 50%.

Table 1: Descriptive Statistics of the Pre-Second Level *Eiken* Test (N=168; k=75)

Min.	Max.	Midpoint	*M*	*SD*
17	64	40.5 (54%)	38.33 (51%)	9.21

The standard deviation of the test scores was 9.21. Figure 1 shows that the scores of 80 Ss (48%) fell below the mean (0-37 points), and 78 Ss (46%) fell above the mean (39-64 points), while the scores of 10 Ss (6%) were exactly on the mean. The results show that the test performed like a true NRT, conforming to a normal distribution.

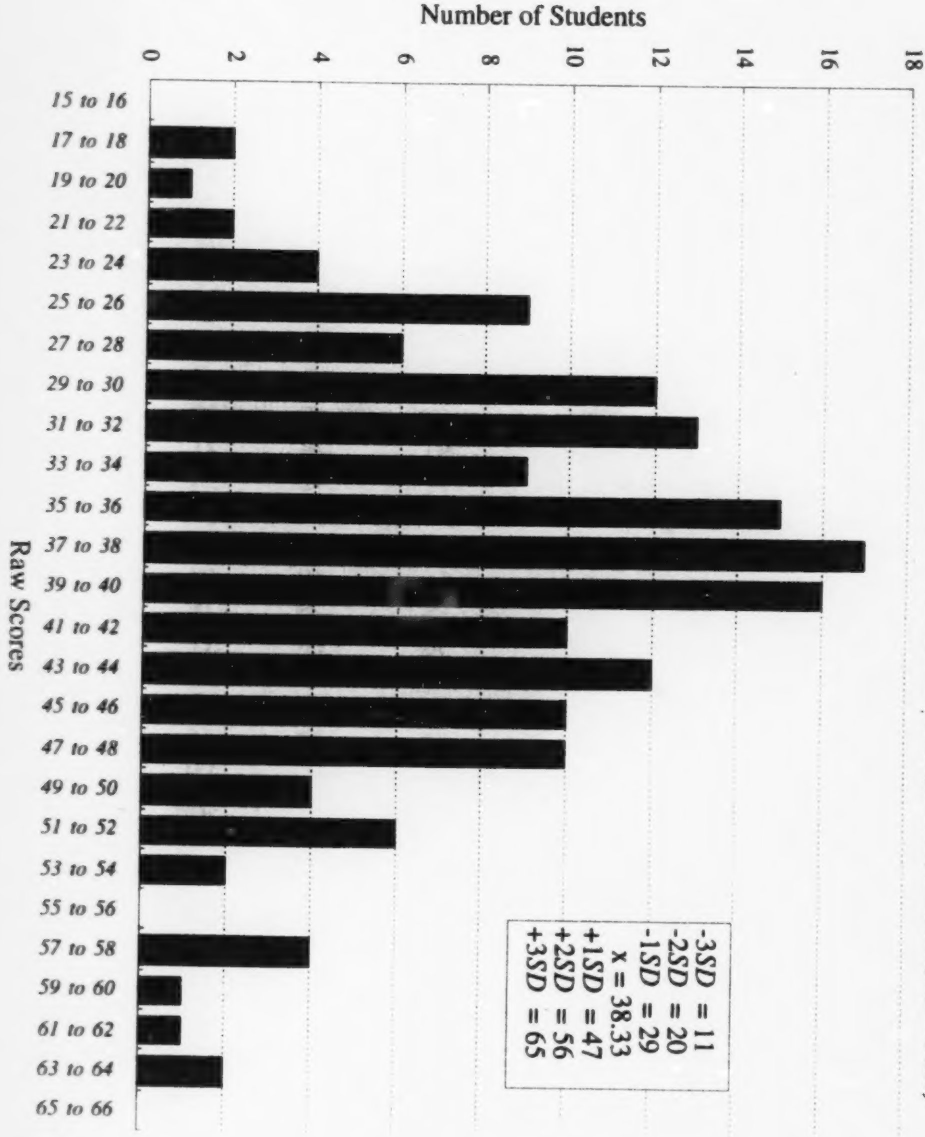


Figure 1: Distribution of Scores on Eiken, Pre-second Level (N=168, x=38.33, *SD* =9.21)

Item Statistics: You will recall that IF shows the percentage of students who answer a given item correctly. For example, 37 Ss answered the first item correctly. Therefore, the IF of item #1 (37/168) is .22 (see Table 2). This means that it was a difficult item because only 22% of the students got it right. Item #3, on the other hand, with an IF of .86, was an easy item for this group.

Table 2: Item Facility (IF) and Item Discrimination (ID) for *Eiken* Pre-Second Grade ($N = 168$)

Item No.	*IF*	*ID*	*Eikyo's IF*†
1	.22	.35	C
2 *	.36	.41	C
3	.86	.17	A
4	.38	.14	C
5	.37	.22	D
6 *	.51	.31	C
7 *	.42	.50	C
8	.26	.08	D
9 *	.50	.42	C
10	.23	.14	D
11	.31	.23	C
12*	.57	.37	B
13	.30	.27	D
14*	.48	.44	B
15*	.58	.42	C
16*	.48	.30	C
17*	.35	.30	C
18	.60	.06	B
19	.12	.17	D
20*	.70	.42	B
21*	.37	.40	C
22	.52	.26	B
23*	.48	.31	B
24*	.41	.48	B
25	.26	.13	D
26	.80	.07	A
27	.35	.23	C
28	.80	.25	A
29	.24	.14	D
30	.52	.21	B
31*	.55	.51	B
32	.40	.20	C
33*	.59	.33	B
34*	.48	.42	B
35	.86	.25	A
36*	.68	.30	B
37	.72	.23	A
38*	.55	.35	B
39	.48	.28	C
40	.25	.25	D
41	.81	.27	A
42	.31	.17	C
43	.72	.44	A
44*	.43	.30	B
45	.82	.27	A
46*	.51	.40	B
47*	.60	.37	B
48*	.55	.44	B
49*	.48	.38	B
50	.28	.21	C
51*	.42	.30	B
52	.38	.22	C
53	.39	.16	B
54	.18	.13	C
55	.28	.16	C
56	.60	.17	B
57*	.66	.31	B
58	.89	.21	A
59	.77	.18	A
60	.86	.09	A
61*	.57	.37	B
62	.39	.02	C
63*	.49	.40	B
64	.65	.16	B
65	.54	.09	C
66	.79	.17	A
67*	.43	.30	B
68	.63	.37	A
69	.23	.07	D
70	.81	.21	A
71	.51	.30	C
72	.94	.03	A
73	.71	.19	B
74	.26	.21	C
75*	.54	.45	B

*= Good test items

† Note: IF reported in *Nihon Eigo Kentei Kyokai*, July, 1996, p. 30.

Item #2 (Table 2), has an IF of .36, and an ID of .41, and therefore meets the above IF and ID criteria. It is fairly well centered and discriminates well between high- and low-scoring students. Other items in the test which could be called good test items are indicated by an asterisk. However, there are many items which discriminate poorly. In fact, more than half of the items (59%) do not meet the ID and IF requirements of good test items, falling outside the acceptable range of .30 to .70. Based on these results, further refinements and improvements of many of the test items are needed.

Eikyo published a table approximating IF by assigning letter values to five ranges (Table 2): A=.8–1.0; B=.6–.79; C=.4–.59; D=.2–.39; E=0–.19 (*Nihon Eigo Kentei Kyokai*, July, 1996, p. 30). Because the ranges are so wide (0.19 points each), however, it is difficult to draw any conclusions about IF from these statistics. Besides, IF alone only tells what fraction of the group got the item correct. In order to draw any concrete conclusions about whether the item is functioning well, ID statistics are needed.

Consistency Estimates: The KR-20 result for this study was .82. It is important to keep in mind that this figure of 82% reliability is based on the in-house scores of this sample of 168 Ss. The actual test population, a much greater number with a broader age range, would have had a different reliability index. Unfortunately, this information has not been made available by *Eikyo*. Only a hint as to the reliability of the *Eiken* test was made by an *Eikyo* representative of the test development section, who said that in the years up to 1992, the reliability of *Eiken* tests was between .80 and .90 (*Eikyo* representative, name withheld, personal communication, July 25, 1995). However, he did not divulge the type of analysis done. In any case, this information is not directly relevant to the present study, which deals with the pre-second level test, first introduced in 1994. It is worth noting that in 1989 and 1990, test reliability for TOEIC using the K-R20 formula was .96 (Woodford, 1992).

The SEM for this test administration was 3.9, meaning that the band around which a student's score should be considered is ±4. Therefore, if a subject who scored 37 on the test were to take the test repeatedly, the scores could vary between 33 (-4) and 41 (+4). The passing level for the June 1996 pre-second level test set by *Eiken* was 38 and above, meaning a subject with 37 would have failed. With no apparent margin for error, all it takes is to be one point short to fail.

Validity

Four general questions pertaining to the test validity were posed. First, is the pre-second level test appropriate for the group *Eikyo* claims?

To determine if the test items were suitable for senior high school level students and above, the items of the June, 1994 pre-second level *Eiken* were compared with the nationally approved senior high school course of study (Wada, 1992) and one Ministry of Education approved textbook, *The Crown English Reading* (Hirano, et al., 1996). This revealed that most of the words and idioms on the pre-second level *Eiken* test are taught at some point during the three years of senior high school. [Exceptions are noted as follows: Section 1, item 10 - *vacant*; item 19 - Bill *came all the way from* Florida; item 20 - *take it easy*; Section 2(B), item 9 - *French, Italian, or Thousand Island salad dressing*; Section 3, item 2 - *Are you having some problems there?*; and Section 4(B), item 9 - *farther inland* (*Nihon Eigo Kentei Kyokai*, 1994b).]

The second question regarding test validity asked whether the contents of the test items reflect aspects of "daily life" in Japan, as *Eikyo* claims (*Nihon Eigo Kentei Kyokai*, 1994a)? At least two do not. The following observations are those of the author, not the results of rigorous evaluations by a team of testing experts.

Section 1, item 10:

When the sign on the door of a rest room says "(   )," it means someone is using it.

1 OCCUPIED   2 VACANT   3 LIMITED   4 EMERGENCY

(*Nihon Eigo Kentei Kyokai*, 1994c, p. 9)

This item is problematic because restroom doors in Japan seldom have signs indicating whether the stall is vacant or occupied. (The exception is on airplanes, where restroom doors are equipped with such signs). The chance that students are familiar with the context of this item, in English or in Japanese, is remote.

Section 2(B), item 10:

A "Do you have the receipt?"
B "Well, it was a present, but it's too small."
C "What's the problem?"
D "I'd like to exchange this skirt, please."

1 C-B-A-D   2 A-D-C-B   3 D-C-B-A   4 B-C-D-A

(*Nihon Eigo Kentei Kyokai*, 1994c, p. 12)

This item, which asks students to put the four sentences in sequential order (3 is the correct choice), is problematic for three reasons: First, it is culturally inappropriate as Japanese do not customarily exchange items of clothing that they have purchased, let alone received as a gift. Second, it is illogical. If the person received the skirt as a gift, it is unlikely that she would have the receipt. The third point is not concerned with item validity as much as consistency. In Section 2(B) there are five items, the first four of which follow a question-answer-question-answer sequence. This item, however, has a statement-question-answer-question structure. The fact that it is different from the other items in this section may be a source of confusion for the test takers. If *Eikyo* intended this confusion, it would be of interest to know the reason.

The third question regarding test validity to be asked is if the items on the test really measure the abilities that they claim to. *Eikyo*'s statement about what the successful pre-second level examinee is able to do, converse, read, and write about daily topics, implies that the *Eiken* tests test all four skills. Only listening and reading are actually tested. The TOEIC is similar to the *Eiken* tests in design, as it too tests only listening and reading. The TOEIC differs from the *Eiken* tests in that it measures listening and reading directly, and speaking and writing indirectly. Validity studies have been done to confirm a high correlation between TOEIC results and speaking and writing skills (Woodford, 1980). As *Eikyo* has not published studies to show correlations between its listening and reading tests to speaking and writing abilities, its claims that the successful examinee is "able to converse about daily topics" and is "able to write about common everyday topics" (*Nihon Eigo Kentei Kyokai*, 1994a) cannot be confirmed.

The final question regarding test validity asked if there were any poorly constructed test items. One example is from Section 3 (item 2):

> Helen: What did you want to talk to me about? You sounded so mysterious on the telephone.
>
> Sherri: Sorry, but I wanted to tell you this news face to face. I've decided to move.
>
> Helen: But I thought you liked your neighborhood. (2)?
>
> Sherri: No, everything is fine. I just need a change.
>
> 2 i) Are you having some problems there? ii) What's the problem?
> iii) Isn't everything fine? iv) Why are you moving?
>
> (*Nihon Eigo Kentei Kyokai*, 1994c, p. 13)

The difficulty with this item is in the first of the four possible responses, which also happens to be the correct answer, "Are you having some problems there?" According to Swan (1995), "*some* is most common in affirmative clauses, while *any* is common in questions and negatives" (p. 548). Further, "we use *some* in questions if we expect people to answer "Yes," or want to encourage them to say "Yes" such as in offers and requests" (Swan, 1995, p. 548) (i.e., *Would you like some more coffee?*). Because choice '(i)' is a question, not an offer or request, it could be argued that it is inappropriate to use *some*.

Two examples of poorly constructed test items occur in the reading passages in Section 4. In the questions following the reading of "Volunteer Guides at Museums" (see Appendix), Item 5 required too much inferencing to make it a viable item:

(5) Professionals at some museums

1 think volunteers should not be paid.
2 feel they know less about museums than volunteers.
3 dislike volunteers because they know more than the professionals.
4 think volunteers cannot do the work of professionals.

(*Nihon Eigo Kentei Kyokai*, 1994c, p. 14)

The correct answer, number 4 (*Nihon Eigo Kentei Kyokai*, 1994c, p. 22), is primarily based on inference, the only clue in the reading passage being, They [professional scholars] feel that amateurs should not do the work of professionals, and that some volunteers act as if they knew everything. The problem is the interpretation of *cannot* in the answer and *should not* in the reading passage.

The second reading passage in Section 4, entitled, "Rainfall in Australia," is problematic in that the text does not correspond to one of the test items (number 7). The pertinent paragraphs and the test item in question are excerpted below:

> Most parts of Australia do not receive enough rainfall. In some places there are long periods when it doesn't rain at all. This lack of rainfall is one of the major reasons why such a large country as Australia has such a small population.
>
> Only one-sixth of the continent—a belt of land along the north, east, and south coasts—receives more than 40 inches of rain a year. The rest receives less than 40 inches, and farther inland are somewhat drier areas that receive between 10 and 20 inches.

(7) Where in Australia do they get more than forty inches of rain?

1 In the center of the south coast.
2 In narrow areas along the coasts.
3 In the areas which have monsoon climates.
4 In wide inland areas.

(*Nihon Eigo Kentei Kyokai*, 1994c, pp. 15-16)

Eikyo states that the correct answer is number 2, which implies that all four coasts receive rainfall. However, the supporting statement in the text is . . . *a belt of land along the north, east, and south coasts—receives more than 40 inches of rain a year*, specifically states only three coasts.

To summarize, it is clear that while most of the vocabulary found on this form of the *Eiken* test is appropriate for the intended examinees, problems of context and item construction make the validity questionable. Without evidence from *Eikyo*, it is difficult to conclude that the *Eiken* pre-second level test is a valid instrument.

How are the Eiken *tests scored?* According to *Eikyo*, the passing score for the pre-second level is "approximately 65%" (*Nihon Eigo Kentei Kyokai*, 1994a, p. 7). However, test score statistics since its introduction in 1994 show that the passing scores are much lower. In 1994, the passing percentages and scores were 55% (41+); in 1995, 56% (42+); and in 1996 (June), 49% (37+). This information is somewhat misleading as it is not made clear that these are standard, not raw, scores.

Eikyo's score reporting system is made an even greater mystery by the fact that students never actually see their test scores. All they receive is a report which states either "pass" or one of three categories of "fail," A, B, C.[5] An *Eikyo* representative explained: "A-level failure encompasses scores up 10 points below the passing score; B is up to 15 points below A, and C covers the remaining scores down to zero" (name withheld, personal correspondence, July 16, 1996). These "guidelines" conflict with a report of the pre-second level test of June, 1994, in which *Eikyo* stated that the passing score was 41; A was 34-40 (7 points below the passing score); B, 27-33 (7 points below A); and C, 26 or lower (*Nihon Eigo Kentei Kyokai*, 1994d). These ranges were consistent for 1995 and 1996 tests as well. The discrepancy between the explanation and the published scoring brings into question the integrity of the reporting system.

The following information (in translation) about how the *Eiken* test is scored was received from an *Eikyo* representative in the Planning Division:

> The passing score for the pre-second level is set at approximately 65%. However, the difficulty of the test inevitably varies from time to time, which leads to adjustment of the passing scores each time a new test is given. In the past, the adjustment of the scores has been done by a thorough item by item analysis, looking at the difficulty of each item [IF], and by using point biserial coefficiency. Recently, *Eikyo* has begun experimenting with another method of analysis, Item Response Theory (IRT), as a replacement for the above-mentioned item analysis procedure. (name withheld, personal correspondence, July 17, 1996)

Point biserial correlation is a calculation which shows item discrimination by computing the correlation between individual item responses and total test scores. Like the ID analysis, an item with a low point biserial coefficient may be discarded from the scoring, resulting in changes in the passing score each time a different test is given. However, there is no indication of whether items have ever been discarded by *Eikyo*.

Item response theory true score equating, also used by TOEFL and TOEIC (TOEFL, 1995, p. 9), converts raw scores to equivalent scaled scores. Although *Eikyo* claims to evaluate its test results using IRT methods, there are no published reports to substantiate these claims.

One final concern regarding test scoring is whether *Eikyo* sets a cutoff for the number of people who can pass. Reports in *Eikyo*'s monthly newsletter, *STEP News*, between 1991 and 1996, and information in their brochure, *The STEP Test?* suggest that this is a possibility. The percentage of people who passed the second level test has been consistent at 18% (1991-1996) and the pre-second level at 30-33% (1994-1996) for a number of years.

Conclusion

The results of this study indicate that the reliability and validity evaluations of the pre-second level *Eiken* test are not favorable. First, the reliability in this study is only .82. Is a test that is 82% reliable good enough? For the uninformed consumer, maybe; for test-makers, definitely not. The validity checks in this study show that the content of the test matches the intended group of test takers, perhaps the test's greatest strength. However, there are problems of clarity and context in the items themselves which need to be corrected. Finally, the item facility (IF) and item discrimination (ID) results in this study indicate that more than half of the test items should be revised or removed as they discriminate only fairly or poorly.

Eikyo has been operating a successful testing business in Japan for more than 30 years. In all likelihood, this trend will continue. However, published reports of studies by *Eikyo* on item construction, reliability, and validity are urgently needed to help consumers become better informed about the test, and to encourage research that would improve the quality of the test so that someday the *Eiken* tests might approach reliability in the high .90s.

Acknowledgments

Special thanks are given to Mikiya Koarai for his assistance with this paper and translations of the communications with Eikyo, *Yuji Ushiro for help deciphering Japanese statistical terms, and two anonymous* JALT Journal *reviewers for their advice.*

Laura MacGregor is an Associate professor at Sapporo International University (formerly Seishu University). She also teaches at Sapporo International University Attached Kindergarten.

Notes

1. The Ministry of Education endorses a total of 15 proficiency tests. In addition to the *Eiken* tests, three others are *Kobitsu Shosha Kentei* (penmanship), *Mohitsu Shosha Kentei* (calligraphy), and *Katei Ryori Gino Kentei* (cooking).
2. The *Eiken* tests are ranked from highest, *i-kyu* (first level), to lowest, *go-kyu* (fifth level). *Kyu* is translated here as level, rather than grade, as more appropriate for the *Eiken* ranking system.
3. The total number of people who took the *Eiken* test from 1990-1994 were as follows: (1990) 2,624,106; (1991) 2,761,771; (1992) 2,830,496; (1993) 2,895,912; and (1994) 3,374,140.
4. There is no second stage interview test for the fourth or fifth levels. For these levels, however the written tests are offered three times a year, in January, June, and October.
5. This style of score reporting is not unique to the *Eiken* tests. It is also used in the tests of secretarial skills (*Hisho Kentei*), and *kanji* proficiency (*Kanji Noryoku Shiken*).

References

Bostwick, M.R. (1995). Evaluating young EFL learners: Problems and solutions. In J.D. Brown & S.O. Yamashita (Eds.), *Language testing in Japan* (pp. 57-65). Tokyo: The Japan Association for Language Teaching.

Brown, J.D. (1995). Differences between norm-referenced and criterion-referenced tests. In J.D. Brown & S.O. Yamashita (Eds.), *Language testing in Japan* (pp. 12-19). Tokyo: The Japan Association for Language Teaching.

Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.

Brown, J.D. & Yamashita, S.O. (Eds.). (1995). *Language testing in Japan*. Tokyo: The Japan Association for Language Teaching.

Gorsuch, G. (1995). Tests, testing companies, educators, and students. *The Language Teacher, 19*(10), 37-41.

Henning, G. (1987). *A guide to language testing*. Boston: Heinle & Heinle.

Hirano, K., Homma, N., Tanabe, Y., Ikegami, Y., Matsuka, H., Yamamoto, S., Sugano, D., Schneider, D., & Wilson, B. (1996). *The crown English reading:*

Teacher's manual. Tokyo: Sanseido.

Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.

MacGregor, L. (1995). More on the *Eiken*. *The Language Teacher, 19*(11), 51, 86.

Nihon Eigo Kentei Kyokai. (1994a). *The STEP test?* Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1994b). *Heisei 6-nen dai 1-kai kentei 1-ji shiken jun 2-kyu (1994, Heisei 6, pre-second level test, 1st stage, 1st administration)*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1994c). *Jun 2-kyu gaido (Guide to pre-second level)*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1994d). *Eiken jun kaijo reporuto (Eiken pre-release report)*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1996, July). *STEP News, 367*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1996, September). *STEP News, 369*. Tokyo: Author.

Swan, M. (1995). *Practical English usage* (2nd ed.). Oxford: Oxford University Press.

TOEFL. (1995). *TOEFL test & score manual, 1995-96 edition*. Princeton, NJ: Educational Testing Service.

TOEIC. (1995). *Guide for TOEIC users*. Princeton, NJ: Educational Testing Service.

Wada, M. (1992). *Kotogakko gakushu shido yoryo (The course of study for senior high school)*. Tokyo: Kairyudo.

Wilson, K. (1993). Relating TOEIC scores to oral proficiency interview ratings. *TOEIC Research Summaries, 1*. Princeton, NJ: Educational Testing Service.

Woodford, P. (1980, November). *The test for English for international communication*. Paper presented at the English-Speaking Union Conference, London.

Woodford, P. (1992). A historical overview of TOEIC and its mission. *The 35th TOEIC seminar* (pp. 10-15). Tokyo: The Institute for International Business Communication.

(Received Sept. 27, 1996; revised Dec. 17, 1996)

Assessing EFL Student Progress in Critical Thinking With the Ennis-Weir Critical Thinking Essay Test[1]

Bruce W. Davidson
Hokusei Gakuen University

Rodney A. Dunham
Tezukayama College

Recent trends in EFL/ESL have emphasized the importance of promoting thinking as an integral part of English language pedagogy; however, empirical research has not established that training in thinking skills can be effectively combined with EFL/ESL instruction. This study made use of the Ennis-Weir Critical Thinking Essay Test to assess progress in critical thinking after a year of intensive academic English instruction among Japanese students ($N = 36$). A control group received only content-based intensive English instruction, while a treatment group received additional training in critical thinking. The treatment group scored significantly higher on the test ($p < .001$). The results imply that critical thinking skills can indeed be taught as part of academic EFL/ESL instruction.

近年、第二言語および外国語としての英語教育では思考力の育成を英語教育の一部として扱うことの重要性が強調されている。しかしながら思考技術のトレーニングが英語教育と効果的に統合できるという結果を示す実証的研究はいまだない。本研究では、Ennis-Weir批判的思考作文テストを用いて、36人の日本人学生が一年間の集中的アカデミックな目的のための英語教育を受けた後にどの程度批判的思考力が発達したかを測定した。統制群は内容中心の英語教育のみを受け、実験群はそれに加えて批判的思考法のトレーニングを受けた。実験群の成績は統制群より有意に高かった。この結果は、批判的思考技術がアカデミックな目的のための英語教育の中で教えることができるのを示している。

Since the advent of research into cognitive development, language teachers and linguists generally have recognized the close connection between language learning and thinking processes. In particular, ESL reading research has shown some correlation between ESL reading comprehension and familiarity with the formal or content schemata of English texts (Carrell, 1987). Furthermore, noting the

JALT Journal, Vol. 19, No. 1, May, 1997

BEST COPY AVAILABLE

unreflective character of many language-teaching approaches that only encourage verbal output or passive input, Tarvin and Al-arishi (1991) and Al-arishi (1994) have explored some methods to make language teaching more thoughtful. Similarly, Chamot (1995) has argued from current educational trends promoting higher-order thinking that EFL/ESL teachers also need to turn the classroom into a community of thinkers. Informal observations may indicate that thinking skills can indeed be taught in an EFL/ESL context (Davidson, 1994, 1995). Without formal testing, however, it is difficult to establish this concretely. Though there has been a lot of thought and research devoted to the development of critical thinking skills in native English speaker educational programs, there has been little research in the area of combining critical thinking with EFL/ESL instruction.

Content-based intensive English instruction has also proven to have many advantages and possibilities (Snow & Brinton, 1988). Is one of them the promotion of critical thinking skills through thought-provoking content? It might be expected that such abilities will develop through discussion, reading, or composition about subjects requiring some serious analytical attention; however, Chance's (1986) survey concluded that critical thinking skills do not develop simply as a by-product of the study of specific subjects. In addition, Halpern (1993) cites evidence from various sources that critical thinking skills can be inculcated through explicit instruction.

These issues and findings inspired a pilot study to discover whether or not critical thinking could be taught to Japanese students of English in a content-based EFL program. After defining what we mean by "critical thinking," we will describe the intensive English program the subjects were enrolled in along with the specifics of the current study. Two research questions guided us:

1. On a critical thinking test task, will English learners exposed to critical thinking skills-training do significantly better than similar students who have not received such training?
2. Can a critical thinking test designed for native English speakers be used as an instrument for evaluating critical thinking skills among non-native English learners?

Critical Thinking: Concept and Inventory of Component Skills

Critical thinking involves rational judgment and discernment of the elements of reasoning. Various definitions of critical thinking reflect this. Norris and Ennis (1989) explain critical thinking as "reasonable and reflective thinking that is focused upon deciding what to believe and do" (p. 3), a definition also stated somewhat differently by Lippman (1991), who defines it as healthy skepticism, and Siegel (1988), who considers the critical thinker to be one who "is appropriately moved by reasons" (p. 2). In contrast to rote memorization or simple information recall, methods for encouraging critical thinking have as their goal the stimulation of the analytical and evaluative processes of the mind (Paul, 1992). Norris and Ennis (1989) have listed a number of critical thinking abilities to develop:

Elementary Clarification
1. Focusing on a question
2. Analyzing arguments
3. Asking and answering questions that clarify and challenge

Basic Support
4. Judging the credibility of a source
5. Making and judging observations

Inference
6. Making and judging deductions
7. Making and judging inductions
8. Making and judging value judgments

Advanced Clarification
9. Defining terms and judging definitions
10. Identifying assumptions

Strategies and Tactics
11. Deciding on an action
12. Interacting with others (p. 14)

This concept of critical thinking and inventory of skills inspired both the instructional treatment and the selection of the Ennis-Weir test in this study.

The Study

Method

Subjects: All participants in the study ($N = 36$) were first-year students enrolled in a private women's junior college in Osaka, Japan. The college's curriculum consisted mainly of an intensive academic English program. Weekly English courses included Oral Discussion (3 hours), Composition (2 hours), Reading (3 hours), Pronunciation (3 hours), and Grammar/Listening (2 hours), totaling 13 hours a week, considering each 50-minute class session as an hour. Oral Discussion, Reading, and Composition followed a topical syllabus of six units of instruction, which included such themes as Prejudice/ Human Rights, Advertising/ Consumerism, and Women's Issues/Child-raising. The integrated, content-based aspect of the program was meant to involve students in in-depth analysis and expression concerning subjects significant in their own lives and in Japanese society. This course of study would seem well-suited to encouraging the development of critical thinking skills as a by-product, since the topics all necessitate thought. Along with the topic, each unit also introduced a rhetorical mode: Illustration, Process, Definition, Classification, Comparison/Contrast, and Persuasion. The first three composition units required students to write a paragraph using each mode, and the last three progressed to multi-paragraph essays. The persuasion essay was written in a mini-term paper format, with references.

In addition to these integrated intensive English courses, Ss took a weekly one-hour seminar course, also conducted in English and concerning some interesting topic or theme such as *American Holidays* or *Traditional Folk Songs*. The treatment group ($n=17$) was composed of students from a seminar on Critical Thinking.[2] Volunteers not enrolled in the critical thinking seminar ($n=19$) served as a control group.

Ss had varying degrees of English proficiency as measured by an in-house proficiency test. At the beginning of the year, this test divided all students into five levels of classes according to scores: A, B, C, D, and E. The A classes had the highest level of proficiency and included students returning from a year or longer of study abroad, whereas the E classes were much less proficient. Regardless of proficiency, however, all classes received similar instruction based on the same content and rhetorical modes noted previously. These Ss represented a broad range of English proficiency levels, as measured by the in-house test, with both groups containing a similar range. Because of the small number of Ss at each level, with the exception of level C ($n=14$), levels were grouped into A+B, C, and D+E for later comparison. The distribution of Ss among the groups was fairly even (control $n=5$, 10, 4; treatment $n=7$, 4, 6). No pre-test was given, in line with the advice of Ennis and Weir (1985), who state that a pre-test is not necessary in research using the test as long as a control group exists. Babbie (1983) has noted that a post-test-only control group design is quite acceptable as long as group assignment is random. Since students enrolled in seminars through a semi-lottery system, the authors consider that in this case group assignment generally embodied the spirit of randomness, although it was not completely random.

Treatment: The treatment group took part in a course designed to train them in basic elements of critical thinking: source credibility, inductive reasoning, informal deductive logic, and assumption-identification. These broad categories encompass most of Norris and Ennis's (1989) list of critical thinking skills, so they were adopted as a framework for the seminar course. During the first semester, instruction dealt with inductive reasoning and source credibility; during the second semester, the emphasis was deductive reasoning and assumption-identification. The course began with an introduction to the concept of critical thinking. Sessions were devoted to exploring various kinds of reasoning fallacies and misuse of evidence, such as over-generalization and the false dilemma (Chaffee, 1990; Damer, 1995). Students were given lists of brief fallacious arguments and asked to explain the problems of each in their own words. In the second half of the semester, the focus shifted to source credibility. Students did exercises in which they evaluated varying accounts of the same event according to differing viewpoints. For example, in groups they discussed and ranked accounts of the results of an international conference, keeping in mind a list of question-criteria: Does the news presenter have a reason to be biased? Is the source an expert in the field? (Beyer, 1991). Students brought in similar examples to present and evaluate.

In the second semester the emphasis shifted to basic argument analysis. First, students did exercises to help them distinguish real arguments from bare claims offering no reason (Engel, 1994). Then they identified the claims and supporting reasons. Later in the semester, the instructor introduced less-obvious aspects of deductive reasoning: unstated assumptions and implications (Scriven, 1976). Using magazine advertisements and other material, students practiced identifying assumptions and implications. As a result of the Hanshin earthquake and other circumstances, a total of only 18 class hours was actually devoted to the course's content.

Instrument: The Ennis-Weir Critical Thinking Essay Test

Test Description: The Ennis-Weir test was chosen for various reasons. One is that it is one of the most generally well-accepted measuring instruments among educators in the critical thinking movement (Walsh & Paul, n.d.), and inter-rater reliabilities have been very high when it has been used (Ennis & Weir, 1985; Hatcher, 1996). Another is that, in contrast to multiple-choice tests, it allows students to justify varying responses, and the test itself presents a realistic critical evaluation task. Other critical thinking tests are available, but almost all of these are multiple-choice instruments that suffer from various weaknesses such as background bias and the impossibility of knowing the reasoning behind an examinee's answer-choice (Ennis, Millman, & Tomko, 1985; Norris & Ennis, 1989). Furthermore, the relatively simple subject-matter and language of the Ennis-Weir test make it suitable for non-native speakers. It has been used successfully with first-year junior high school native English speakers in the U.S.

The test itself contains a simple set of instructions and a letter to a newspaper editor containing 10 brief paragraphs. The fictional writer, Raywift, recommends that overnight parking be prohibited on all the streets of his town, Moorburg. After a brief introduction, eight numbered paragraphs elaborate the argument. Most are weak and commit various common reasoning fallacies such as equivocation, irrelevancy, poor statistical sampling, and circular reasoning, but some contain legitimate support, consisting in the use of qualified experts or a relevant reason. Point-by-point, the examinee's task is to judge the thinking of each of these numbered paragraphs and to evaluate the strength of the letter's argument as a whole in a final summary-paragraph. For example, in responding to Paragraph 3, examinees are expected to notice that a relevant reason is offered to support Raywift's argument. Similarly, test-takers are supposed to show some indication that they comprehend the flaws of the experiment in Paragraph 6. (See Appendix for copies of the introduction and paragraphs 3, 6, and 8.)

A clear and specific scoring protocol accompanies the test indicating various possible answers and how each is to be scored. Points are awarded both for judging correctly and for indicating a valid reason for

Table 1: Critical Thinking Skills Addressed on the Ennis-Weir Test

Paragraph	Skill
1.	Noticing misuse of analogy and/or shift in meaning
2.	Recognizing irrelevant reasoning
3.*	Recognizing relevant reasoning
4.	Recognizing circularity and/or the lack of a reason
5.	Recognizing defective reasoning
6.	Recognizing insufficient sampling
7.	Recognizing equivocation and/or the use of an arbitrary definition
8.*	Evaluating the credibility of expert testimony

*Paragraphs that exhibit sound reasoning one's judgment, and a penalty point of -1 can be deducted for poor reasoning. Each answer can receive a maximum of 3 points and a minimum of -1, except for the summary paragraph, where a maximum of 5 points can be awarded. Therefore, the overall score can range from -9 to +29. In general, the protocol gives latitude to raters to award points whenever an examinee can give a credible reason in support of this evaluative judgment, even when it differs from that of the protocol writers. Brief answers are acceptable as long as they indicate a valid judgment, backed up with a sound reason for that judgment.

We considered the possibility that cultural differences between America and Japan might bias the results, but we felt that in this case culture would not be a significant issue. For one thing, the Moorburg letter concerns street parking, which is not a point of significant cultural difference between the U.S. and Japan. Street-parking laws also exist in Japan and are stringently enforced. If anything, parking is more of a problem in Japan, so Japanese readers might be even more likely to identify with a parking-related issue than Americans. Furthermore, writing a letter to an editor complaining about a public problem, as in the Raywift letter, is a common practice in Japan as well. The directness and abrasiveness of Raywift's style are perhaps the only aspects of the letter that might seem strange or unsettling to a Japanese reader. However, it should also be noted that abrasive political rhetoric is not unknown in Japan.

A limited amount of research has been done in the U.S. using the test. The largest study to date has been Hatcher's (1996) at Baker University. Over a period of four years (1990-1994), American freshmen scored an average of 11.8 to 13.8 on the Ennis-Weir test after a year-long compulsory critical thinking course. They had scored from 5.8 to 9.4 on a pretest and registered gains of 2.8, 5.8, 5.8, and 6.0 points. Interestingly, a number of Chinese and Japanese students at Baker University also took part in the study, but their scores were eliminated from it because they consistently scored poorly. Hatcher (1994) speculated that their low scores may be due to Oriental politeness and accordingly a hesitancy to criticize the Moorburg letter.

Test Administration: In the last week of second semester classes, the Ennis-Weir Critical Thinking Essay Test (1985) was administered, with the control group given the test within the same week. Both groups had 80 minutes to read the test and write nine brief paragraphs in response, twice the amount of time recommended by the test-makers. Since the subjects were non-native English speakers, it was felt that more time would be necessary for them to comprehend the material and compose answers. To help them with the language aspects of the test, they were allowed to use dictionaries. Furthermore, before taking the test, all subjects received two sample test items with model answers to make sure the Ss understood two things: (a) that they had to make a clear evaluative judgment as to whether the argument in each paragraph was a good one or not and (b) that they had to give a clear reason or explanation for their judgment. Without such explicit direction, the subjects might not have done either of these two things. However, students in the present study were used to doing peer-evaluation of essays in composition classes, so the idea of writing comments or criticism about a piece of writing was already somewhat familiar. Basic information about the Ss, including English level and overseas experiences, was collected on the answer form.

Results

Tests were scored blindly and independently by two raters. The test-raters in this study found little difficulty in using the protocol to judge student answers. Grammatical or vocabulary problems were overlooked unless they made an answer incomprehensible. Inter-rater reliability was found to be adequate ($r = .72$). The scores and information collected were examined. Therefore, the average scores given to each student for each of the 10 scores on the test, one score for each of the nine paragraphs and a total score, were used for all subsequent analysis.

The small number of Ss in the study (control $n = 19$ and treatment $n = 17$) makes relationship detection difficult unless it is very strong. Because this is the first study of this type, we were interested in detecting moderate relationships as well as strong ones. Consequently, we decided that the risk of committing a Type I error would be less important than missing moderate relationships. Therefore, the significance level of .10 was chosen as the cutoff for accepting or rejecting relationships. Nevertheless, we have reported here the exact probability for all results that indicated statistically significant relationships.

The most important analysis, of course, dealt with the effect of critical thinking training on test scores. Therefore, we conducted a t test to compare the scores of the two groups. The treatment group's mean score on the Ennis-Weir Test was statistically 6.6, significantly higher than the control group's mean score of 0.6 (t (27.73) = -4.99, $p < .001$). Table 2 shows the range of scores for each group and details the differences. As the table shows, 10 Ss in the treatment group scored 7 or

Table 2: Group Comparative Scores on the Ennis-Weir Test

Score Range	Control Group	Treatment Group
-4.0 to 0.0	9	1
1.0 to 2.0	6	1
2.5 to 6.5	4	5
7.0 to 13.5	0	10
Total N	19	17
Mean	0.6	6.6
Median	1.0	7.5
Mode	-1.5	3.0

higher, while only one scored 0 or lower. In contrast, in the control group, nine Ss scored 0 or lower, with no score higher than 6.5.

Next, the individual paragraph scores of the control and treatment groups were compared. There were statistically significant differences between the mean scores of the control group and the treatment group on two paragraphs: the third paragraph (M=-0.18 and 0.50, respectively, t (33.87)=-2.10, p=.043) and the sixth paragraph (M=-0.55 and 0.15, respectively, t (27.10)=-2.59, p=.015). The difference in scores on the eighth paragraph approached statistical significance (t (33.69)=-1.71, p=.096), with the treatment group scoring higher (M=0.47) than the control group (M=-0.52). However, the scores on the remaining paragraphs showed no statistically significant difference between the control and treatment groups (p>.10).

Since the test was in English, a foreign language for the Ss, proficiency may have affected scores. As mentioned, since the number of Ss at each proficiency level, except C (n=14), was quite small (A level n=7, B level n=5, D level n=5, and E level n=5), A level was grouped with B, and D level with E. The C level was left intact, creating a three-level variable. The distribution for this ordinal variable was compared for levels A+B, C, and D+E between the control group (n=5, 10, 4) and the treatment group (n=7, 4, and 6). As determined by a chi square test, the distribution indicates that there was no relationship between English proficiency and the type of group ($\chi^2(2, N=36)=3.204$, p=.202). An analysis of variance was also run to examine the relationship between English level and test scores. There was no statistically significant relationship between the two variables ($F(2, 35)=1.57$, p=.224). Table 3 shows that the range of scores was comparable for each proficiency level. Judging by this analysis as well as the phrasing of student answers on the test, we believe that students generally did not do poorly simply as a result of an inability to understand the English of the test. The sample test items appeared to succeed in helping students to grasp the kind of test task they were engaged in, and the wording of the test did not appear to present an insurmountable problem even for lower-level Ss.

Table 3: Scores on Ennis-Weir Test by English Level

Score Range	A-B Ss	C Ss	D-E Ss
-4.0 to 0.0	4	4	2
1.0 to 6.5	4	8	4
7.0 to 13.5	4	2	4
Total *N*	12	14	10
Mean	4.3	1.8	4.7
Median	3.8	1.5	4.3
Mode	10.5	1.0	3.0

The next variable examined was overseas experience, since a number of Ss had lived a year or longer in an English-speaking country. Using a t test, scores of Ss who had traveled overseas and those who had not were compared. The differences in scores between the two groups was not significant for total scores or for any individual paragraph score ($p > .10$).

Because the Ennis-Weir test deals with parking problems, each of the subjects was asked to report on the test form whether or not she possessed a driver's license. Japanese students often first learn to drive at the age of 19 or 20, the age of the Ss in this study, and familiarity with driving an automobile may have helped some Ss do better on the test, which concerns a parking problem. Scores for the two groups, those with licenses and those without, were compared. Total scores were statistically the same for both groups; however, students without driver's licenses ($M = -0.66$) scored statistically significantly lower on the seventh paragraph than Ss with ($M = -0.17$) [$t\,(21.14) = -1.84$, $p = .079$]. On the eighth paragraph, those without licenses ($M = -0.24$) also scored significantly lower than Ss with ($M = 0.70$) [$t\,(31.50) = -3.31$, $p = .002$]. Otherwise, the fact of having a driver's license showed no significant relationship with student scores ($p > .10$). Since the specific issues addressed in the seventh and eighth paragraphs are not directly related to the experience of driving we consider the statistical significance to be unrelated to the current study.

These statistical analyses appear to indicate that the differences in scores between the treatment and control groups cannot be accounted for by differences in English proficiency levels or other factors such as overseas experience or having a driver's license. Therefore, the differences in scores on the Ennis-Weir Test can probably be attributed to the critical thinking training given the treatment group.

Discussion and Conclusion

Both research questions can be answered in the affirmative, based on the results of this study. The Ennis-Weir test, designed for native English speakers, appears to be usable for non-native English learners. Furthermore, it is encouraging to find that even a small amount of instruction in the basics of critical thinking appeared to result in higher scores for the treatment group. Critical thinking skills can apparently be taught to some extent along with English as a foreign language and can, therefore, enhance a content-based course of study. In view of the relatively small amount of actual instruction, the rather low average score of 6.6 is not surprising, and is much better than the performance of the control group ($M=0.6$). As a point of comparison, Baker University American freshmen registered gains of 2.8, 5.8, 5.8, and 6.0 (Hatcher, 1996) in four successive years. Interestingly, three of those gains approximate the difference of 6.0 that we found in the mean scores of our two groups, though the mean score of the treatment group (6.6) is only half that of the average post-test scores of the Baker freshmen. Looking at the individual test items, differences between the two groups appeared specifically in items which had received instructional attention in the critical thinking class. Paragraph 6 deals with the misuse of statistics, a reasoning problem dealt with in class, while Paragraph 3 featured a relevant reason, another instructional point. The difference in performance on Paragraph 8, which concerned the use of experts and their credibility as sources, also approached statistical significance, and that area also had received attention in the source-credibility component of the critical thinking seminar. In contrast, little difference in scores appeared in the case of Paragraphs 1 and 7, which both concerned inappropriate definitions, an area not dealt with in the course. Furthermore, there was little difference in scores on Paragraph 4, which consists in circular reasoning and was very similar to one of the sample test items. Perhaps because of its similarity, 35 Ss responded correctly to it.

The overall quality of the answers of the two groups differed, but they shared certain tendencies indicating a general weakness in the area of critical thinking skills. This is not surprising in view of the fact that Japanese education does not seem to encourage debate or the critical evaluation of reasoning (Davidson, 1995). Detailed consideration of the answers themselves is beyond the scope of this study, but it is revealing to explore the kinds of errors consistently made by the participants. All of the subjects had been taught to identify and use definition, illustration, and argumentation as rhetorical modes; however, this training apparently did not prepare them to recognize reasoning errors related to these modes. For example, in the case of the 24 Ss who positively but incorrectly evaluated Paragraph 2, all gave as their justification the fact that Raywift provided a reason grounded in reality or else that he gave a concrete example. They missed the fact that both the reason and the example were irrelevant to his argument. Similarly, students accepted the definition-arguments in Paragraphs 1 and 7, even though the definitions offered by Raywift were inappropriate. For instance, he argues in Paragraph 7 that his opponents "don't know what 'safe' really means. *Conditions are not safe if there's even the slightest possible chance for an accident"* [italics added] (Ennis & Weir, 1985, p. 13). Only 2 Ss found fault with this impractical definition of the concept of safety; the others credited him with giving a clear definition. Likewise, 25 of 36 Ss accepted the false analogy used in Paragraph 1. Though the treatment group fared better on some paragraphs and in their overall scores, these common tendencies seem to point to a general need for critical thinking training among these particular Japanese EFL students that perhaps is not being addressed adequately by practice in English rhetorical modes or content-based study. It is even possible that exposure to rhetorical modes such as definition, illustration, and argumentation may only predispose students to accepting weak ideas simply because they are presented in the proper rhetorical format. Without concurrent attention to reasoning fallacies and the pitfalls related to each mode, teachers may discover that for their EFL/ESL students, a little bit of knowledge of rhetorical modes is a dangerous thing. Such students may one day find themselves struggling with the reasoning tasks required in an English academic setting, regardless of their general English language proficiency or familiarity with English modes of expression.

Though Hatcher (1994) speculated that politeness and a hesitancy to make negative judgments may have inhibited Japanese and Chinese performance on the Ennis-Weir test at Baker University, answers in our study reveal that they did not err only in positively evaluating weak paragraphs. They also often negatively assessed Raywift's better arguments in Paragraphs 3 and 8. For example, a number of Ss rejected

Raywift's citation of qualified experts in Paragraph 8 as "just opinions." Ss did not appear to suffer from any hesitancy to criticize.

This is only a limited pilot study, and more research of a similar type needs to be done to substantiate these tentative conclusions. Larger student samples are needed. Also, it would be helpful if a translated version of the test could be administered to groups of similar Japanese students to remove completely the possibility that English language deficiencies may to some extent account for the lower scores. For cultural and linguistic reasons, however, such a translated test may be difficult to make and administer. Students in mixed-nationality EFL/ESL programs in other cultural settings could also provide interesting and relevant data about critical thinking abilities and the possibility of developing and testing them in English language programs, since English language-learning problems related to thinking are not confined to Japan. English instructors in other places have noted reasoning weaknesses similar to the ones we have found (Sherman, 1992; Matthews, 1994). Furthermore, it would be informative to experiment with other standard tests of critical thinking in EFL/ESL programs. Finally, it is worth exploring the question of whether training in critical thinking can improve general English language proficiency, especially in writing and reading. Nevertheless, we hope to see the Ennis-Weir test applied by others in studies bearing some similarity to ours. This relatively unexplored area invites further inquiry.

Bruce W. Davidson teaches English at Hokusei Gakuen University in the Department of Social Welfare and researches issues related to critical thinking education. Send correspondence to: z00134@hokusei.ac.jp.

Rodney A. Dunham teaches in the English Department at Tezukayama College and conducts research in mass communication.

Notes

1. Earlier versions of this paper were presented November 5, 1995 at the Twenty-First Annual JALT (Japan Association for Language Teaching) International Conference, Nagoya, and July 29, 1996 at the Sixteenth International Conference on Critical Thinking and Educational Reform, Rohnert Park, CA.
2. The treatment group consisted of 17 out of 22 members from the seminar on critical thinking. Five could not take the Ennis-Weir test due to circumstances arising from the Kobe-Osaka earthquake.

References

Al-arishi, A. (1994). Practical ways to promote reflection in the EFL/ESL classroom. *English Teaching Forum, 32*(2), 2-5.

Babbie, E. (1983). *The practice of social research* (3rd ed.). Belmont, CA: Wadsworth.

Beyer, B. (1991). Practical strategies for the direct teaching of thinking skills. In A. Costa (Ed.), *Developing minds: A resource book for teaching thinking: Vol. 1* (2nd ed., pp. 274-279). Alexandria, VA: Association for Supervision and Curriculum Development.

Carrell, P. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly, 21*, 461-481.

Chaffee, J. (1990). *Thinking critically*. New York: Houghton-Mifflin.

Chamot, A. (1995). Creating a community of thinkers in the ESL/EFL classroom. *TESOL Matters, 5*(5), 1, 16.

Chance (1986). *Thinking in the classroom: A survey of programs*. New York: Teachers College Press, Columbia University.

Damer, E. (1995). *Attacking faulty reasoning: A practical guide to fallacy-free arguments*. Belmont: Wadsworth.

Davidson, B. (1994). Critical thinking: A perspective and prescriptions for language teachers. *The Language Teacher, 18*(4), 20-26.

Davidson, B. (1995). Critical thinking education faces the challenge of Japan. *Inquiry: Critical Thinking Across the Disciplines, 14*(3), 41-53.

Engel, M. (1994). *With good reason: An introduction to informal fallacies*. New York: St. Martin's Press.

Ennis, R., Millman, J., & Tomko, T. (1985). *Cornell Critical Thinking Tests Level X & level Z* [manual]. Pacific Grove, CA: Critical Thinking Press and Software.

Ennis, R., & Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test*. Pacific Grove, CA: Critical Thinking Press and Software.

Halpern, D. (1993). Assessing the effectiveness of critical-thinking instruction. *JGE: The Journal of General Education, 42*(4), 238-254.

Hatcher, D. (1994, August). *Grading the Ennis-Weir Critical Thinking Essay Test*. Paper presented at the Fourteenth International Conference on Critical Thinking and Educational Reform, Rohnert Park, CA.

Hatcher, D. (1996). Combining critical thinking and written composition: The whole is greater than the sum of the parts. *Inquiry: Critical Thinking Across the Disciplines, 15*(2), 20-36.

Lippman, M. (1991). *Thinking in education*. Cambridge: Cambridge University Press.

Matthews, C. (1994, March). *Using news stories to develop thinking/reasoning skills*. Paper presented at the TESOL annual convention, Baltimore.

Norris, S. & Ennis, R. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Critical Thinking Press and Software.

Paul, R. W. (1992). *Critical thinking: What every person needs to survive in a rapidly changing world*. Santa Rosa, CA: Foundation for Critical Thinking.

Scriven, M. (1976). *Reasoning*. New York: McGraw-Hill.

Sherman, J. (1992). Your own thoughts in your own words. *ELT Journal, 46*(2), 190-198.

Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and educa-*

tion. New York: Routledge Press.
Snow, M. & Brinton, D. (1988). Content-based language instruction: Investigating the effectiveness of the adjunct model. *TESOL Quarterly, 22*, 533-574.
Tarvin, W., & Al-arishi, A. (1991). Rethinking communicative language teaching: Reflection and the EFL classroom. *TESOL Quarterly, 25*, 9-27.
Walsh, D., & Paul, R. (n.d.). *The goal of critical thinking: From educational ideal to educational reality*. American Federation of Teachers Educational Issues Department.

(Received July 3, 1996; revised Nov. 20, 1996)

Appendix: The Moorburg Letter (Introduction and paragraphs 3, 6, and 8)

Dear Editor:

Overnight parking on all streets in Moorburg should be eliminated. To achieve this goal, parking should be prohibited from 2 a.m. to 6 a.m. There are a number of reasons why any intelligent citizen should agree.

3. Traffic on some streets is also bad in the morning when factory workers are on their way to the 6 a.m. shift. If there were no cars parked on these streets between 2 a.m. and 6 a.m., then there would be more room for this traffic.

6. Last month, the Chief of Police, Burgess Jones, ran an experiment which proves that parking should be prohibited from 2 a.m. to 6 a.m. On one of our busiest streets, Marquand Avenue, he placed experimental signs for one day. The signs prohibited parking from 2 a.m. to 6 a.m. During the four-hour period, there was *not one accident* [italics added] on Marquand. Everyone knows, of course, that there have been over four hundred accidents on Marquand during the past year.

8. Finally, let me point out that the Director of the National Traffic Safety Council, Kenneth O. Taylor, has strongly recommended that overnight street parking be prevented on busy streets in cities the size of Moorburg. The National Association of Police Chiefs has made the same recommendation. Both suggest that prohibiting parking from 2 a.m. to 6 a.m. is the best way to prevent overnight parking.

Sincerely,
Robert R. Raywift

(Ennis & Weir, 1985, p. 13)

Contrastive Rhetoric in Letter Writing: The Interaction of Linguistic Proficiency and Cultural Awareness

Taeko Kamimura
Senshu University

Kyoko Oi
Toyo Gakuen University

This study examines the relationship between Japanese college students' proficiency in English and their cultural awareness toward a target-language culture (North American). Japanese EFL students were divided into four groups according to high and low English proficiency and high and low cultural awareness based on performance on the CELT and the researchers' cultural awareness (CA) test. The four groups wrote college and scholarship application letters. The letters were analyzed in terms of frequency and content of different Semantic Formulas. The results show that subjects with both high English proficiency and high cultural awareness manifested the rhetorical patterns closest to those in the native speakers' English letters of application. Subjects with low English proficiency and low cultural awareness showed the rhetorical patterns closest to those in the Japanese letters of application. Subjects who lacked either the sufficient level of English proficiency or cultural awareness produced writing which varied from the target style. Results indicate that cultural awareness may be as important an element as English ability in student writing.

本研究は日本の大学生の英語能力と目標言語（アメリカ英語）の文化に対する気づきの間の関係を考察するものである。日本の大学生4グループが大学入学申込および奨学金申請の手紙を書いた。被験者は、CELTによって測定された英語能力と筆者の作成した文化的気づきを測定するテストの結果に従ってグループ分けされた。手紙は意味公式の頻度と内容という観点から分析された。その結果、英語能力が高く文化的気づきも高い被験者は母語話者の書く同様の手紙に最も近い修辞パターンをしめした。英語能力も文化的気づきも低い被験者は、日本語で同様の目的の手紙を書く場合に最も近い修辞パターンを示した。英語能力または文化的気づきのどちらかが欠ける被験者は、目標のスタイルとは異なる手紙を書いた。これらの結果は、文化的気づきが学生の英作文能力の重要な要因であるかもしれないことを示している。

JALT Journal, Vol. 19, No. 1, May, 1997

BEST COPY AVAILABLE

Robert Kaplan first proposed the notion of contrastive rhetoric in 1966. Since then many researchers of writing across the world have engaged in active research in this field. Enough evidence has been reported to support Kaplan's claim that each language has a culturally-preferred way of organizing ideas in discourse (e.g., a linear development in English vs. an indirect approach in a gyre style in Oriental languages), and that writers from different linguistic-cultural backgrounds transfer their preferred discourse patterns when they write in other languages. Among those researchers who have dealt with the differences of Japanese and English are Hinds (1979, 1980, 1981a, 1981b, 1987), Connor and McCagg (1983, 1987), Kobayashi (1984), Oi (1986), Mok (1993), Fisher-Stoga (1993, 1995), and Kimball (1996). All of them, however, have dealt with expository or argumentative writings in academic settings. Research oriented toward more pragmatic, non-academic perspectives has been scarce; only Jenkins and Hinds (1987) discusses the rhetorical differences in a more pragmatic context, namely business letter writing.The present study involves the comparison of rhetoric in letter writing in English and Japanese. This study is enforced by the following observations: 1) letter writing (such as social letters and business letters) is important to Japanese EFL students since it is the type of English writing that the students will face most frequently; 2) among several kinds of letter writing, however, the one they face immediately would be a letter of application because of an increasing number of Japanese students wish to study in American colleges, and 3) letter writing carries a pragmatic function to convey an intended meaning to a specific target audience in a particular culture.

The scheme of the present study is derived from the findings of two previous studies. The first study (Oi & Sato, 1990) investigated whether rhetorical transfer would be observed in Japanese EFL students' letter writing, comparing their letters with those of native speakers of English. The comparison was threefold. The first group was composed of Japanese students writing in Japanese, to determine the nature of Japanese rhetoric; the second group was composed of Japanese students writing in English; and the third group was native speakers of English writing in English. They all wrote under the same directions with the same assignment, though the directions were given in different languages depending on the groups: i.e., the first group in Japanese and the second and third groups in English. In this research, not only did we find rhetorical differences in letter writing between Japanese and American writers, we also found noticeable evidence of rhetorical transfer by Japanese students. However, this research did not take into consideration the English proficiency levels of Japanese students, so whether or not the difference was merely due to developmental factors was unknown.

The second study (Sato & Oi, 1990) was conducted under a similar format. However, this time the Japanese students were divided into two groups, high and low, according to level of English proficiency. The study showed rhetorical transfer was observed across the proficiency levels, indicating that the presence or absence of students' rhetorical transfer is not determined by English proficiency level alone. That is, there must be other factors.

The Study

From our previous studies we found that English proficiency alone is not a decisive factor in affecting Japanese EFL students' writing behavior in letter writing. Therefore, in the present research, we introduced a new factor, cultural awareness, which is in the domain of pragmatic competence. We define cultural awareness as one's familiarity with the perceptive and behavioral patterns in a target culture, American culture in this study. Combining these two factors, we propose the following research questions:

1) Does the degree of cultural awareness affect the Japanese EFL students' writing behavior in letter writing?

2) If the degree of cultural awareness is related to the students' writing behavior, what are the roles of cultural awareness and English proficiency respectively in the students' letter writing?

Method

Subjects: Subjects ($N=42$) were selected from students enrolled in three college EFL writing classes. The Ss were sophomore English majors at a Japanese college. They had practiced some narrative and expository writings, but they had experienced no formal training in letter writing. The Ss included in the study were selected, following the procedure outlined below.

Procedure

Task: All students in the three classes were asked to write a letter of application to a college in English. They were told to read a notice which announced the offering of a scholarship at a college in America and to write a letter applying for the college and the scholarship.

Instruments: Based on the results of the two former studies, we introduced two new factors for consideration. One is the variation in English proficiency levels of Japanese EFL students, and the other is the degree of their cultural awareness of American culture.

In order to measure the English proficiency levels, we administered the CELT (Comprehensive English Language Test) (Harris & Palmer, 1986). To test the degree of cultural awareness, we devised our own cultural awareness (CA) test. For the cultural awareness test, we selected 20 questions regarding critical situations that reflect crucial differences between Japanese culture and American culture (see Appendix). These questions covered the social, school, and workplace environments. The questions were devised by us, but the ideas were extracted from various books about cross-cultural communication that emphasize the differences between Japanese culture and American culture (e.g., Condon & Yousef, 1975; Sakamoto & Naotsuka, 1982; Barnlund, 1975, 1989; Furuta, 1987, 1990; Nishida, H., 1989; Nishida, T., Nishida, H., Tsuda, & Mizuta, 1989; Sherard, 1989; Matsumoto, 1994). The CELT and CA were administered to all students in the three classes. Based on the scores, we divided the students into the following four groups, W, X, Y and Z (Table 1). The W Group included students who scored high in both the CELT and CA test; the X Group scored high in the CELT but low in the CA test; the Y Group scored low in the CELT but high in the CA test; the Z Group scored low on the both tests. The demarcation line for "high" was placed at the upper 30 percent line among all the students and "low" at the lower 30 percent line for both tests. That was 165 points for high and 140 points for low on the CELT, and 15 points for high and 11 points for low on the 20-point CA test. Students who fell into these categories were selected as the subjects (N=42). Other students were excluded from the study.

Table 1: Classification of Subjects

CELT	CA High ≥ 15	 Low < 11
High ≥ 165	W (n = 9)	X (n = 9)
Low < 140	Y (n = 8)	Z (n = 16)

W: CELT-High, CA-High; X: CELT-High, CA-Low; Y: CELT-Low, CA-High; Z: CELT-Low, CA-Low

Analysis

Each of the letters written by the Ss was segmented according to idea units called Semantic Formulas (SFs), as done in our previous two studies. This analytical measure was based on a study by Beebe, Takahashi, and Uliss-Weltz (1990), which dealt with the analysis of refusals in speech. The SFs employed in the present study and a sample of each SF are:

1. Identification (ID): *I am a student at — college.*
2. Social talk (ST): *How are you?*
3. Referring to the ad. (RE): *I saw your ad about the scholarship.*
4. Writing a letter (WR): *So I am writing an application letter.*
5. Application message (AP): *I decided to apply.*
6. Reason (REA): *I'm interested in American culture and to study in America.*
7. Qualification (QUA): *I have a 3.8 grade point average (on a 4.0 scale) at —, and a score in the upper 20% bracket on the SAT test.*
8. Disqualification (DIS): *I'm afraid of going to the U.S. by myself.*
9. Petition (PE): *Could you please help me?*
10. Personal appeal (PA): *With these experiences, I feel I could make a positive contribution to ABC College and hope you will consider my application.*
11. Reference (REF): *I am enclosing a reference from Mr. Kempski, Head of the History department.*
12. Promise (PR): *I'll study hard.*
13. Apology (AP): *I'm sorry I have a favor.*
14. Request for information (REQ): *Please send me any forms that need completing.*
15. Closing remark (CR): *I would appreciate for your kindness.*
16. Expecting a reply (EX): *I'm looking forward to your letter.*

We first analyzed each letter into a sequence of the different SFs (see "Analysis of Sample Writing" below for sample analyses). We further analyzed those SFs in two aspects: 1) frequency (how often they appeared) and 2) content (concrete examples of SFs). The results were compared with those obtained in the two earlier studies in respect to the nature of mother-tongue writing of Japanese colleges students writing in Japanese and the target-language writing of native speakers writing in English.

Results and Discussion

Frequency

Each application letter was segmented according to the 16 different SFs, and how often each SF appeared was examined across the four different groups. Table 2 shows the frequencies of the 16 SFs identified in the four groups as well as those identified in the American subjects writing in English and the Japanese subjects writing in Japanese in the Oi & Sato study (1990).

In previous studies (Oi & Sato, 1990; Sato & Oi, 1990), the following differences were found between the Japanese and American subjects' application letters:

Table 2: Frequency of the Semantic Formulas Used by Different Groups

SF	Group					
	American ($n=13$)	W ($n=9$)	X ($n=9$)	Y ($n=8$)	Z ($n=16$)	Japanese ($n=30$)
ID	11 (84.6%)	8 (88.9%)	6 (66.7%)	5 (62.5%)	10 (62.5%)	16 (53.3%)
ST	0 (0%)	0 (0%)	2 (22.2%)	1 (12.5%)	9 (56.3%)	8 (26.7%)
AD	6 (46.2%)	5 (55.6%)	4 (44.4%)	4 (50%)	8 (50%)	25 (83.3%)
WR	6 (46.2%)	0 (0%)	1 (11.1%)	1 (12.5%)	2 (12.5%)	6 (20%)
AM	10 (76.9%)	9 (100%)	6 (66.7%)	6 (75%)	11 (68.8%)	18 (60%)
REA	6 (46.2%)	9 (100%)	7 (77.8%)	7 (87.5%)	13 (81.3%)	30 (100%)
QUA	8 (61.5%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
DIS	0 (0%)	0 (0%)	2 (22.2%)	1 (12.5%)	3 (18.8%)	5 (16.7%)
PE	1 (7.7%)	0 (0%)	3 (33.3%)	2 (25%)	8 (50%)	13 (43.3%)
PA	5 (38.5%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
REF	3 (23.1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
PR	0 (0%)	1 (11.1%)	2 (22.2%)	3 (37.5%)	6 (37.5%)	13 (43.3%)
AP	0 (0%)	0 (0%)	2 (22.2%)	0 (0%)	0 (0%)	0 (0%)
IM	7 (53.8%)	5 (55.6%)	2 (22.2%)	2 (25%)	5 (31.3%)	7 (23.3%)
CR	7 (53.8%)	3 (33.3%)	1 (11.1%)	0 (0%)	2 (12.5%)	14 (46.7%)
EX	2 (15.4%)	3 (33.3%)	2 (22.2%)	2 (25%)	3 (18.8%)	2 (6.7%)

ID=Identification; ST=Social talk; RE=Referring to the ad; WR=Writing a letter; AP=Apllication message; REA=Reason; QUA=Qualifier; DIS=Disqualification; PE=Petition; PA=Personal appeal; REF=Reference; PR=Promise; AP=Apology; REQ=Request; CR=Closing remark; EX= Expecting a reply

1) Many American subjects used *persuasive* strategies, showing their abilities and previous experiences with concrete data. They often referred to their academic records and underscored their qualifications.

2) The Japanese subjects, regardless of the language used (English or Japanese) or of their English proficiency levels, tended to use *emotional* strategies meant to attract the reader's sympathy.

The Japanese subjects realized these *emotional* strategies by using such SFs as "social talk," "disqualification," "petition," and "promise." Those four SFs were rarely found in the letters by the American subjects. Our previous studies suggested that these four formulas were evidence of the Japanese subjects' negative transfer from Japanese writing and that these formulas were culturally inappropriate in English letter writing when addressed to American readers.

The SF "social talk" is a clear transfer of the Japanese usual letter format in which one is supposed to open a letter message with either greetings such as *I trust this finds you in good health*, or references to the weather such as *It has been awfully hot this summer. How are you coping with the heat?* It is considered impolite just to start business abruptly in any Japanese letter, even in a business letter. The Japanese students are so much used to this writing convention that they tend to persistently transfer this SF negatively in English application letters.

The SF "petition" is a desperate-sounding plea such as *Please give me a chance*. This type of pathetic tone is unfitting in an English letter of application.

The SF "disqualification" is a statement like *My English is not good*. Judging from the American practice of a letter of application to a college, this statement is like a taboo, admitting a reason for disqualification as an applicant. We interpret this as a reflection of the Japanese tendency to value modesty or understatement of oneself as often described in many studies (among them, Condon & Yousef, 1975). Concerning the Japanese tendency to resort to "petition" and "disqualification" when writing a letter of application, an exemplary anecdote is introduced in Sakamoto and Naotsuka (1982, pp. 34-37).

A typical example of the SF "promise" is a statement such as *I'll study hard if I'm admitted to your school*. A number of Japanese students finish their letters with "promise" following "petition." We regard it as a reflection of the Japanese formulaic expression *gambarimasu*, which can be translated as *I'll do my best*.

The present study, therefore, focuses on those four SFs (social talk, disqualification, petition, and promise) particularly identified as typical in Japanese Ss' letters of application.

Table 3 shows the frequencies (in percentage) of the four SFs used by the different groups (W, X, Y, Z) as well as those used by the American subjects writing in English and the Japanese subjects writing in Japanese (Oi & Sato, 1990).

Table 3: Frequencies of the Four Sematic Formulas (%)

SF (Sematic Formulas)			Group			
	American	W	X	Y	Z	Japanese
ST (Social talk)	0	0	22.2	12.5	56.3	26.7
DIS (Disqualification)	0	0	22.2	12.5	18.8	16.7
PE (Petition)	7.7	0	33.3	25	50	43.3
PR (Promise)	0	11.1	22.2	37.5	37.5	43.3

Each of the four groups demonstrated unique writing patterns. The writing patterns of the W Group, high English proficiency and high cultural awareness, were most similar to the patterns found for American subjects in the previous studies. No subjects in the W Group employed "social talk," "disqualification," or "petition," and only one subject included a "promise" (11.1%).

In contrast, the Z Group, low English and low cultural awareness, varied most from the native English speakers' pattern and showed the clearest transfer from Japanese. The Z Group ranked highest in the frequencies of three SFs: "social talk" (56.3%), "disqualification" (18.8%), and "petition" (50%), and second highest in the frequency of "promise" (37.5%).

The X Group, high English proficiency and low cultural awareness, and the Y Group, low English proficiency and high cultural awareness, were between the W and Z Groups. Though the differences between the X and Y Groups were not large, the percentages of the three out of four Japanese-oriented SFs were higher in the X Group (22.2% for "social talk," 22.2% for "disqualification," and 33.3 % for "petition") than in the Y Group (12.5% for "social talk," 12.5% for "disqualification," and 25% for "petition"). This suggests that both the writing patterns of X and Y Groups differed less from the target pattern than the Z Group, but the two groups did not as closely approximate the target style as the W Group. Additionally, the X Group demonstrated a greater variation from the target pattern than the Y Group.

Content

The next step of analysis was to examine what the subjects actually wrote for each of the four SFs in the letters of application. For each SF, several subcategories were established to examine the actual content of the four SFs in detail. For example, the SF "social talk" was further classified into such subcategories as *Hello, How do you?* and *How are you?* Table 4 illustrates the results of the content analysis.

The content written by the Z Group varied most from the pattern for the English letter of application and reflected most clearly the practice of the Japanese letter of application. The typical pattern of a letter of application by the Z Group started with a colloquial social talk *Hello* (31.3%), said, *I don't have enough English ability* (12.5%), yet pleaded for the scholarship by saying, *Please give me the scholarship* (25%), and ended with a promise by saying, *I will study hard* (31.3%). Thus, the letters by the Z Group had emotional and pathetic tones. In contrast, there was only one example of "promise" for the W Group: *I will make efforts* (11.1%). Emotional and pathetic tones were not perceived in the letters by the W Group.

The X and Y Groups manifested only one or two examples of some subcategories of the four SFs. The general tendency, however, was for both the X and Y Groups to be positioned between the W and Z Groups, as was found in the results in the frequency count of the SFs. The letters by the X and Y Groups did not sound as pathetic and emotional as those by the Z Group, but they were not as completely free of these tones as those by the W Group. Furthermore, such emotional and pathetic tones were somewhat stronger in the letters of the X Group than those of the Y Group.

Analysis of Sample Writing

Representative samples of the letters done by Ss from each of the four groups, W, X, Y, and Z, help explain the characteristic writing patterns. Each sample letter is analyzed by SFs, with errors left intact.

Sample 1, written by S-1 in the W Group (high English proficiency and high cultural awareness), is quite close to the target letters of application by the native speakers of English. The letter concisely conveys the intended message by including such SFs as "identification," "reason," "application message," "request for information," and "closing remark." None of the four SFs which characterize Japanese letters of application (social talk, disqualification, petition, and promise) are included.

Table 4: Content of the Four Semantic Formulas

SF Subcategory	Group					
	American(*n*=13)	W (*n*=9)	X (*n*=9)	Y (*n*=8)	Z (*n*=16)	Japanese(*n*=30)
Social Talk						
Hello	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	5 (31.3%)	2 (6.7%)
How do you do?	0 (0%)	0 (0%)	1 (11.1%)	1 (12.5%)	2 (12.5%)	6 (20%)
How are you?	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (12.5%)	0 (0%)
Disqualification						
Don't know reality	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
Don't have enough English ability	0 (0%)	0 (0%)	1 (22.2%)	1 (12.5%)	2 (12.5%)	4 (13.3%)
Afraid of going to America	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
Cannot express opinions	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	1 (6.3%)	0 (0%)
Petition						
Help me	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	2 (6.7%)
Admit me	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (18.8%)	2 (6.7%)
Understand me	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	7 (23.3%)
Give me the scholarship	0 (0%)	0 (0%)	1 (11.1%)	1 (12.5%)	4 (25%)	1 (3.3%)
Give me a good answer	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
Give me a chance	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (6.3%)	0 (0%)
Promise						
Study hard	0 (0%)	0 (0%)	1 (11.1%)	2 (25%)	5 (31.3%)	5 (16.7%)
Make efforts	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
Lead a full life	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (6.7%)
Make good use of the scholarship	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
Get something in America	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	0 (0%)
Do my best	0 (0%)	0 (0%)	0 (0%)	1 (12.5%)	0 (0%)	0 (0%)
Mature	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (6.3%)	0 (0%)

Sample 1:

S-1, W Group, High English Proficiency and High Cultural Awareness

Dear Mr. Thompson:

***IDENTIFICATION**

[I am a student in _________ University in Japan and I'm very interested in your exchange program between ABC College and our university.]

***REASON**

[I'm studying English and American Literature here and I believe studying in your college will much help my research in as well as improving my English skills.]

***APPLICATION MESSAGE**

[Therefore I do want to apply for this program. And I would like to apply for the scholarship you offer simultaneously.]

***REQUEST FOR INFORMATION**

[Please send me a brochure and/or more details about them.]

***CLOSING REMARK**

[Your most up-to-date information will be tremendously appreciated.]

Sincerely yours,

Sample 2 was written by S-2 in the X Group (high English proficiency and low cultural awareness). Since the level of S-2's English proficiency is high, there are no major grammatical mistakes. It does not include such Japanese-oriented SFs as "petition" and "promise." However, S-2 employs "disqualification" and underestimates self-worth, in contrast with the practice in the English letter of application where the use of "persuasive" strategies is expected. For instance, S-2 writes, *I can neither understand what English speakers say nor express my thought in English well*, and *I'm not good at express my opinion to other people*. Thus, even though S-2 exhibits no substantial grammatical mistake in the domain in linguistic competence, problems in cultural awareness in the realm of pragmatic competence are exhibited.

Sample 2:
S-2, X Group, High English Proficiency and Low Cultural Awareness

•REASON

[I have studied English since I was a junior high school student. English is not so easy to learn, but I'm interested in it very much, because the pronunciation of English words is very different from one of Japanese words, and I like pronouncing them very much.]

•DISQUALIFICATION

[To my regret, Japanese education in English is not so good for learning English conversation. I don't think I have much trouble reading English, but I can neither understand what English speakers say nor express my thought in English well.]

•REASON

[Living in the United States is the best way to improve such troubles of mine. That is the first reason I want to study at your college. The second one is that I want to take part in an active lessons.]

•DISQUALIFICATION

[I'm not good at express my opinion to other people.]

•REASON

[It' OK in Japan, but it cannot be allowed in other countries, so I want to train myself in active discussions in lessons at a college in the United States.

I have been to the United States to learn English before. My parents paid for me in that case. I thank them very much, but I can't have them pay any more.]

•APPLICATION MESSAGE

[So I'd like to be offered scholarship.]

Sample 3, from S-3, is representative of the Y Group (low English proficiency and high cultural awareness). S-3 does not include any of the four Japanese-oriented SFs. However, S-3's lack of organization makes the style far different from that found in the target letters of application. Reasons for application are mentioned sporadically and the self-introduction is begun, in an inappropriate place, too abruptly.

Like Sample 2, Sample 3 is not an appropriate English letter of application, but for different reasons. Although Sample 3 has no critical problem in content from a pragmatic aspect, it has many problems in grammar and organization.

Sample 3:
Y Group, S-3, Low English Proficiency and High Cultural Awareness

Dear Mr. Thompson,

* APPLICATION MESSAGE

[I want to get the scholarship.]

*REASON

[Because My father has been sick since last year. And My family is very poor. I can't afford money to go the college.]

*APPLICATION MESSAGE

[But I want to study English and literature in ABC college. I want to go to America.]

*REASON

[I'm interested in American costums, culture, family life and eating life. And I want to understand American people and watch beauties of nature, town.]

*IDENTIFICATION

[Introduction my self and my family.
My name is __________. I'm nineteen years old. I'm _________ University college student. My hobbies are playing tennis, watching movies, cooking, and shopping. I have four members. My father, My mother, my brother, me.
My father is 57 and businessman. But he is sick now. My mother is 48 and House wife. My brother is 23 and he graduated University this spring but he doesn't catch a job and in house he is studying law everyday for exercise.]

*REASON

[I want to learn Literature and American life. So I want to speak English.]

Sample 4, by S-4, represents the Z Group (low English proficiency and low cultural awareness). Sample 4 is linguistically unacceptable, with many errors in sentence construction and no organization as a paragraph. S-4 includes three of the four Japanese-oriented SFs: "social talk," "promise," and "petition." The letter begins with *hello*, makes several promises, saying, *If I go to ABC College, I study harder than now*, and *I will grow than now and I will come back to Japan!* In addition, S-4 petitions Mr. Thompson, saying, *Mr. Thompson, I want to know American people and culture. Please get the chance to me*. Thus, it is inappropriate as an English letter of application, both linguistically and culturally.

Sample 4:
S-4, Z Group, Low English Proficiency and Low Cultural Awareness

Dear Mr. Thompson:

***SOCIAL TALK *IDENTIFICATION**

[Hello,] [My name is ____________. I am twenty years old now. I am interested in American people and culture. But I've never seen foreign countries. I want to go to American very much. Of course, I am studying hard very day.]

***APPLICATION MESSAGE *REASON**

[I want to get the scholarship.] [Because to help my home's life. My brother is high school student and my home is very new. Going to America need much money.]

***PROMISE**

[If I go to ABC College, I study harder than now.]

***REASON**

[And I want to make many foreign friends there. I think American is very friendly and kindly. Sure, I will get nice relationship with them.]

***PETITION**

[Mr. Thompson, I want to know American people and culture. Please get the chance to me.]

***PROMISE**

[After year I will grow than now and I will come back to Japan!]

Conclusion

In this study, we have examined the rhetorical differences between Japanese and English letter writing. The Japanese EFL students in this study seem to transfer Japanese rhetorical patterns into English when they write in English. Of immediate concern to the present study was examination of two factors, English proficiency and cultural awareness, which may determine the degree of rhetorical transfer.

The first research question sought to determine whether the degree of cultural awareness affects Japanese EFL students' letter writing behavior. The results of the present study answer this question affirmatively: both the level of English proficiency and the degree of cultural awareness of the target culture affect Japanese EFL students' letter writing behavior. The second research question concerned the respective roles of cultural awareness and English proficiency in the Japanese EFL students' writing performance. The results show the following tendencies:

1. Students with high English proficiency and high cultural awareness produce letters closest in style to that of native speakers of English;
2. Students with low English proficiency and low cultural awareness produce letters closest in style to that of native speakers of Japanese;
3. Students with high English proficiency and low cultural awareness produce letters with culturally inappropriate content but acceptable English;
4. Students with low English proficiency and high cultural awareness produce letters with generally culturally appropriate content but problematic English.

In order for our EFL students to compose in a way which is acceptable to an English-speaking audience, we need to develop not only their English proficiency but also their cultural awareness. This is especially so in letter writing, which carries a more pragmatic function than writing such as exposition and argumentation. As Mok (1993) asserts:

> Awareness of the [cross-cultural] differences is important because it makes students realize that to become part of the target language discourse community, they need to develop new attitudes, to meet certain criteria of the target language's traditions, and, in some cases, to put aside their native language habits. (p. 157)

We need, therefore, to develop teaching methods and teaching materials which integrate cultural factors with linguistic ones. Questions such as those used in our cultural awareness test could be modified and turned into instructional tasks. Alternatively, students could conduct their own SF analyses: first writing application letters, then analyzing the SFs in the draft to see whether L1-based SFs were included, and finally revising the letters into acceptable English letters of application.

More research will be needed to determine whether cultural awareness is a critical factor when the Japanese EFL students engage in other types of letter writing. In addition, research will be needed to examine whether the types of cultural and rhetorical instruction suggested above have positive effects on EFL students' writing performance.

Acknowledgments

We are grateful to the editor and two anonymous JALT Journal *reviewers for their careful reading and valuable comments.*

Taeko Kamimura teaches EFL writing and applied linguistics at Senshu University. Her interests include the relationship between L1 and L2 composing processes.

Kyoko Oi teachers EFL writing and applied linguistics at Toyo Gakuen University. Her research interests include the relationship between language and culture.

References

Barnlund, D. C. (1975). *Public and private self in Japan and the United States*. Tokyo: Simul Press.

Barnlund, D. C. (1989). *Communicative styles of Japanese and Americans*. Belmont, CA: Wadsworth Publishing Company.

Beebe, L, Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R. C. Scarcella, E S. Anderson, & S. D. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 55-73). New York: Newbury House.

Condon, J. C., & Yousef, F. (1975). *An introduction to intercultural communication*. New York: Macmillan Publishing Company.

Connor, U., & McCagg, P. (1983). Cross-cultural differences and perceived quality in written paragraphs of English expository prose. *Applied Linguistics, 41*, 259-268.

Connor, U., & McCagg, P. (1987). A contrastive study of English expository prose paragraphs. In U. Connor, & R. B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 73-86). Reading, MA: Addison-Wesley.

Fisher-Stoga, F. (1993). Convention and composition in the Japanese ki-sho-ten-ketsu: Towards a methodology of contrastive rhetoric. *Tokyo Daigaku Gaikokugoka Kenkyukiyo (The Proceedings of the Department of Foreign Languages and Literature, College of Arts and Sciences, University of Tokyo), 41*(3), 130-168.

Fisher-Stoga, F. (1995). An independent learning: Lafcadio Hearn as educator. *Tokyo Daigaku Gaikokugoka Kenkyukiyo (The Proceedings of the Department of Foreign Languages and Literature, College of Arts and Sciences, University of Tokyo), 43*(3), 24-43.

Furuta, G. (Ed.) (1987). *Ibunka komyunikeshon (Intercultural communication)*. Tokyo: Yuhikaku.

Furuta, G. (Ed.) (1990). *Ibunka komyunikeshon kii waado (Key words for intercultural communication)*. Tokyo: Yuhikaku.

Harris, D. P., & Palmer, L. A. (1986). *CELT (A Comprehensive English Language Text for learners of English)*. New York: McGraw-Hill Book Company.

Hinds, J. (1979). Organizational patterns in discourse. In T. Givon (Ed.), *Syntax and semantics 12: Discourse and syntax* (pp. 135-157). New York: Academic Press.

Hinds, J. (1980). Japanese conversation, discourse structure, and ellipsis. *Discourse Processes, 3*, 263-286.

Hinds, J. (1981a). *Ellipsis in Japanese*. Alberta: Linguistic Research.

Hinds, J. (1981b). Japanese expository prose. *Papers in Linguistics, 13*, 117-157.

Hinds, J. (1987). Reader versus writer responsibility: A new typology. In U. Connor & R. B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 141-152). Reading, MA: Addison-Wesley.

Jenkins, S., & Hinds, J. (1987). Business letter writing: English, French, and Japanese. *TESOL Quarterly, 21*, 327-345.

Kaplan, R. (1966). Cultural thought patterns in inter-cultural education. *Language Learning, 16*, 1-20.

Kimball, J. (1996). Writing on academic topics: Externalizing rhetorical processes in an intercultural context. *JALT Journal, 18*, 55-66.

Kobayashi, H. (1984). Rhetorical patterns in English and Japanese. *TESOL Quarterly, 18*, 737-738.

Matsumoto, S. (1994). *Nihonbunka no tokushitsu (Characteristics of Japanese culture)*. Tokyo: Kenkyusha.

Mok, W. E. (1993). Contrastive rhetoric and the Japanese writer of EFL. *JALT Journal, 15*, 151-161.

Nishida, H. (1989). *Jitsurei de miru nichibei communication gap (Communication gap between Japan and America examined by examples)*. Tokyo: Taishukan.

Nishida, T., Nishida, H., Tsuda, Y., Mizuta, S. (1989). *Kokusai ningen kankeiron (Studies on international human relationships)*. Tokyo: Seibunsha.

Oi, K. (1986). Cross-cultural differences in rhetorical pattering: A study of Japanese and English. *JACET Bulletin, 17*, 23-48.

Oi, K., & Sato, T. (1990). Cross-cultural rhetorical differences in letter writing: Refusal letter and application letter. *JACET Bulletin, 21*, 117-136.

Sakamoto, N., & Naotsuka, R. (1982). *Polite fictions*. Tokyo: Kinseido.

Sato, T., & Oi, K. (1990). *Nichiei retoriku no hikaku bunseki* (A contrastive analysis of Japanese and English Rhetoric), In JAAL-in-JACET (Ed.), *Oyoogengogaku no kenkyu (Studies in applied linguistics)* (pp. 146-155). Tokyo: Liber Press.

Sherard, M. (1989). *Tacit assumptions*. Tokyo: Macmillan Language House.

(Received March 1, 1996; revised September 2, 1996)

Appendix: Sample Questions from Cultural Awareness Test

Under the following circumstances, which behavior do you think you are likely to follow? Choose one from the two alternatives.

[Social life]

1. Six months have passed since you came to the U.S. with your parents. Today you have invited Mr. and Mrs. Brown to your home. They have been very kind to you since you came to the U.S. and this is their first visit. While chatting over tea, Mrs. Brown says to you, "Could I see the rest of your house?" To your regret, the house is far from being clean enough to show to other people. How would you react to Mrs. Brown's request?
   1) You would refuse, saying, "It's such a mess. I really cannot show you this time."
   2) You would show her around, saying, "It's a mess, but if you don't mind that."

2. You are asked to have tea by an American woman, Mrs. Anderson. It is hot, and you are thirsty. When Mrs. Anderson says, "What would you like? Would you like something hot or cold?" what would you say to her?
   1) You don't want to bother Mrs. Anderson, so you would say, "Anything will be fine with me."
   2) You would say clearly what you would like to have.

[School life]

3. You are now studying at D University in the U.S. and taking Sociology I among other courses. Professor Samson, who is teaching Sociology I, takes a discussion style in his class. Since your English is still not good enough, you cannot quite participate in the discussion with American students, although you are trying to. There is another Japanese student, Mariko, in this class. She is always quiet and does not contribute to the discussion. You came here two years earlier than Mariko. How would you advise her?
   1) You would suggest that she tell Professor Samson of her linguistic disadvantage and ask him to acknowledge her willingness to participate.
   2) You would advise her to participate in the discussion as actively and assertively as she can, seeking the professor's help after class as needed.

4. It has been a month since you began studying at B University in the U.S. The other day you were asked to give a speech for an audience comprised of professors. Although you are not confident of your English as it has been just a month, you've decided to give the speech. How would you deliver the speech?
   1) You would try to be confident of your English and not mention anything concerning the ability of your English.
   2) You would, first of all, tell the audience that your English is not good because you are afraid that the audience will be surprised at your poor English.

[Workplace]

5. You are employed by an American company. Yesterday you saw Jane, who is a co-worker, step into the elevator before Mr. Black, who is her boss. You are older than Jane. How would you feel about her behavior?

1) You would assume it natural since she is a woman.
2) You would try to reprimand her as you think she was being rude.

6. After graduation from college, you climbed up the ladder of success and are now a branch-office manager. As business is good this year, you are quite busy. Today you have work that needs to be done by tomorrow. Unfortunately tomorrow is Sunday. If you fail with this, it means a loss to the company, so you want your employees to come to work tomorrow. What would you do?

1) You would ask your employees to come to work even on Sunday, explaining to them it is for the sake of the company.
2) You would ask for volunteers to help with the project, stressing extra benefits for those who choose to do so.

実際使用アクティビティーを使用した海外での日本語教育例

加藤好崇

東海大学留学生センター

ネウストプニー（１９９５）によるとインターアクション教育のためには、「解釈アクティビティー」、「練習アクティビティー」、「実際使用アクティビティー」の三つのアクティビティーの適用が必要である。まず、「解釈アクティビティー」で講義などによる説明で知識を与え、「練習アクティビティー」でロールプレーなどの形で練習をする。そして、最後に「実際使用アクティビティー」において実際使用場面を教育過程の中に導入する。本稿は、上記の考え方をハバロフスク国立教育大学日本語学科において適用した授業報告例である。ここでは、日本人ビジネスマンへのインタビュー依頼の手紙、電話での場所や日時等の決定、インタビュー、結果発表などが「実際使用アクティビティー」として導入されている。また、それぞれの段階において「解釈」と「練習」のアクティビティーも「実際使用アクティビティー」に先立って実施された。最後にこのプロジェクトワークの結果得られた言語学上、社会言語学上、社会文化上の有効点、問題点を考察する。

Neustupny (1995) claims that learners need to be engaged in three kinds of activities to acquire interactive competence in a second language. The first of these is interpretation activities. These aim to provide learners with linguistic knowledge, including explanations by lectures and the learners' own research in the form of project work. Second is practice activities, which entail drills, role plays, and simulations. Third is performance activities. These differ from practice activities in that they provide authentic situations for learns to interact in the target language. Examples of performance activities are immersion courses and inviting native speakers of the TL to the classroom. This article reports the process and outcome of project work organized by two classes of advanced learners of Japanese at a Russian university following the framework suggested by Neustupny. This project focused on opinions of Japanese native speakers (NSs) working in Khabarovsk about Japanese and Russian working customs. The project, conducted during 10 classes of 80-minute each, followed five stages: preliminary preparation stage, writing a letter asking for an interview, making an appointment on the phone, interviewing a Japanese NS, and presentation of the results to the entire class. Each stage consisted of a sequence of interpretation activities, practice activities, and performance activities.

BEST COPY AVAILABLE

JALT Journal, Vol. 19, No. 1, May, 1997

Interviews were recorded and used for follow-up interviews with the subjects. In addition, the author interviewed four of the eight Japanese NSs who had been interviewed by the subjects. From the analyses of the recorded NS/subject interviews, the follow-up interviews with subjects and NSs, and the content of the presentations, it was found that: 1) some of idioms and vocabulary encountered by the subjects for the first time during their interviews of the NSs were adopted into the subjects' active vocabulary; 2) some of the vocabulary provided by the instructor in the preliminary stage was reinforced through use in the interviews and presentations, again entering the subjects' active vocabulary; 3) subjects became aware of the variety of Japanese dialects; 4) subjects became aware of the Japanese view toward Russian customs, and 5) subjects participated more actively and spontaneously during the project than they did during classroom-based activities. Major problems perceived were: 1) the use of Japanese dialects by the NSs made comprehension difficult for the subjects; 2) the interviews were not always well structured and some of the Japanese NSs interviewed felt that they had been asked an unorganized list of questions, and 3) the content and the language of the answers by the Japanese NSs were not adjusted to the linguistic level of the subjects, who were then unable to use appropriate communication strategies to repair communication breakdowns. The author suggests similar projects be undertaken, especially ones with follow-up interviews of all the NSs interviewed, in order to establish clear criteria for teaching and evaluation. The author also suggests teacher intervention during the planning stage, especially in the preparation of interview questions, to aid learners in organizing the interview and keeping the language from the NSs at a level the learners would find manageable.

海外で日本語を学ぶ学生にとって、教師以外の日本人とのコミュニケーションは非常に貴重な体験となる。またその際、情報の伝達を目的としていく過程で、学生の目標言語を使用する能力も促進されると考えられる（Ellis, 1990)。本稿ではこうした考え方に基づき、筆者が勤務したロシア連邦ハバロフスク国立教育大学において試みられた授業例を紹介したい。この授業では、在住日本人へのインタビュー依頼の手紙、電話でのアポイントメント、実際のインタビューなどが主要なタスクとして組み込まれた。

この授業の具体的な目的は以下の３点である。

① 学生に日本人との実際のインターアクションを体験させ、またその中で日本語使用能力の向上を計る。

② 学生の学習意欲を高める。

③ 学生に日本語の中に存在する方言などの多様性や日本人の考え方の多様性を認識させる。

タスクベースシラバス

タスクの捉え方は研究者によって多少の差があるが（Richards, Platt & Weber, 1986; Nunan, 1989)、基本的には、目標言語を理解あるいは産出する過程で何らかの課題を達成していくことを言う。そして、このタスクによって授業はよりコミュニカティブになる可能性を持つ。

Nunan (1989, pp. 59-63)はこのタスクを構成する要素の一つである活動を次のような観点から分類している。第一に、活動自体のauthenticityの程度、第二に、その活動がスキルを身につけることを目的としているものか、実際のコミュニケーションの中でスキルを使うことを目的としているものか。第三に、それが流暢さを目的としているものか、正確さを目的としているものか、の三点である。しかし、これらの活動が行なわれる環境に関しては、Nunanは主に教室内に中心を置いている。

それに対して、Strevens (1987)は教室外で行われる活動の重要性を指摘している。活動例としては、ペンパルを持つこと、目標言語話者のビジネスコミュニティーと関係を持つこと、メディアを利用するものなどが挙げられている。またStrevensは、教室外で行われる活動は学生が実際のインターアクションに教師を介さず接することができ、また活動を管理する責任が学生に移動することにより学習意欲が高まるなどの利点を挙げている。

ネウストプニー(1995)もまた教室外での活動の必要性を認め、それを「実際使用アクティビティー」と呼ばれる活動の範疇に含めている。ネウストプニー(1995, p.23)によると、「実際使用アクティビティー」はその前段階として、知識を与えることを目的とした「解釈アクティビティー」、次にその知識を使って練習を繰り返す「練習アクティビティー」の存在が必要であり、実際のインターアクションを目的とする「実際使用アクティビティー」は最後に行なわれなければならないとしている。ネウストプニー(1995, pp. 20-30)はこの三つの「アクティビティー」について以下のように説明している。

①「解釈アクティビティー」: 最も伝統的、かつ広く行われている方法で、具体的には講義による説明や学習者がプロジェクトの形などで「調べ」たりするものなどが含まれる。

②「練習アクティビティー」: インターアクションの練習のためのものである。このアクティビティーには、文法的ドリル、ロールプレー、シミュレーションなどが含まれる。しかし、これらはあくまで練習として行われるもので自然さの程度に問題が残り、さまざまな状況が予想される実際の場ですぐに効果があるかどうか疑問である。

③「実際使用アクティビティー」:「練習アクティビティー」と違い、実際の目標言語の使用場面を教育現場に取り入れたものである。例えば、教室内にお客を呼んで学生に質問をさせたり、あるいは学校内のいろいろな場所を案内させる、日本料理店に行くなどの活動が考えられる。また、オーストラリアのモナシュ大学で行われているイマーションコース（尾崎・ネウストプニー,1986）などもその例として挙げらる。

本稿の授業例も、この三つの「アクティビティー」の流れに沿って実施された。

BEST COPY AVAILABLE

実施コースの概要

本稿で述べるインタビューを主とした一連の授業は1994年11月から12月までの10数回（行事などの都合でクラス間に多少の差が生じた）の授業（各8 0 分）で実施された。実施された学年は3年（12名）および4年（12名）である。このプロジェクトでは3名ずつが1つのグループを形成し行動をとった。

テーマは各学年とも「ハバロフスクで働く日本人（日本人の労働観）」であり、これは筆者によって選択された。実施コース全体の構成は図1のようになる。

図1：実施コースの構成

①「予備段階」　テーマに沿った予備的な知識を与える
↓
②「手紙を書く段階」*<1解釈→2練習→3実際使用>インタビューの依頼の手紙を書く
↓
③「電話でアポイントメントをとる段階」<1解釈→2練習→3実際使用>
インタビューの日時を決定する
↓
④「インタビュー段階」<1解釈→2練習→3実際使用>約束場所に行き、インタビューを実施
↓
⑤「発表段階」<1解釈→2練習→3実際使用>　インタビューの結果をグループ毎に発表

*「解釈」は「解釈アクティビティー」、「練習」は「練習アクティビティー」、「実際使用」は「実際使用アクティビティー」を表わす。

また、実際のインタビューの会話は学生によって録音され、筆者がそれを文字化した後、発表内容との比較および分析を行った。さらに、筆者はプロジェクト終了後、各グループのメンバーをそれぞれ呼び、インタビューの録音を聞きながら、フォローアップインタビュー（ネウストプニー，1981，pp. 36-38）を試みた。また、インタビューを受けた日本人のうち、4名にもフォローアップインタビューを行った。

最後に、評価に関しては、今回は「発表段階」に限り、デス・マスの一貫、発表内容の明瞭さ・掘り下げ方・面白さ、敬語、明瞭化への配慮、発音などの項目で評価を行った。

各段階の詳細

「予備段階」（3，4回分の授業）

この段階で学生はテーマに関連する背景的知識を学習した。内容としては、現在の日本の労働状況、女性の労働も含めた問題点、就職状況、サラリーマンの日常などが取り上げられ、以下の教材を中心としてすすめられた。またこの際、聞き取り問題、読解問題などのタスクが学生に課せられた。

＜教材：①ビデオテープ「サラリーマンの休日」（国際交流基金）②「日本を話そ

う」（ジャパンタイムズ）③「総合日本語中級」（凡人社）④「朝日新聞で日本を読む」（くろしお出版）⑤朝日新聞などから関連記事を利用した生教材＞

次に上で得た知識をロシア人や個人の立場で考え直させるために、ディベート（3対3で、「人生には余暇が最も大切」グループ対「人生には仕事が最も大切」グループ）とレポート（ロシア人の労働観や労働状況、あるいは日本との比較についてで、これは宿題として提出させた）を課した。

また、この段階までに筆者はインタビューを受けてくれる在住日本人を探し出し、インタビューの許可を得ている。そして、各グループにインタビューの相手を自由選択させた。

「手紙を書く段階」

①「解釈アクティビティー」（1回半～2回分の授業）

ここでは、「手紙の書き方事典」（三省堂）を基礎とし、手紙の基本的な構成、特殊な言い回しの例、あるいは手紙のサンプルを見せるなどして説明を行った。今回は特に未見の人に出すインタビュー依頼の手紙なので、それに関連する項目を中心に説明した。

②「練習アクティビティー」（一回分程度の授業、宿題）

グループごとにインタビュー依頼の手紙を作成。出来上がったものを黒板に書かせ、筆者が教室内でチェックした。この際、手紙の構成や特別な表現、敬語に特に注意させた。宿題として、筆者が日本に帰って一年経過したという想定で日本の筆者に出す手紙の作成をさせた。

③「実際使用アクティビティー」

学生は教室内で訂正された「依頼の手紙」を宿題として書き直しをした後、提出。筆者の許可がおり次第、清書し郵送した。具体的な手紙の内容としては、自己紹介、事情説明とインタビューの趣旨及び依頼、後日電話にてインタビューを受けてもらえるかどうか確認をしたい、といったものが含まれている。次のページの例はその一例である。

「電話でアポイントメントをとる段階」（1，2回分の授業）

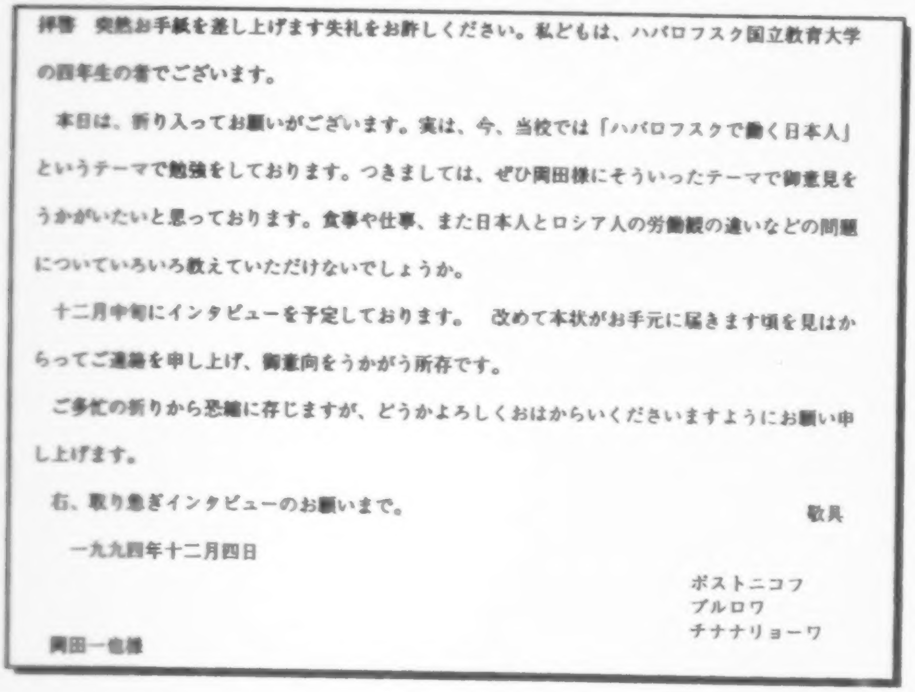
①「解釈アクティビティー」

まず、各グループごとに、電話で話すべき内容（自己紹介、インタビューの目的、インタビューを受けてもらえるかどうかの確認、日時、場所などを）を確認させ、その流れを簡単にロシア語でメモさせる。次に一人代表者を出し、筆者と電話でアポイントメントをとるロールプレーをし、それを録音する。この内容を全体で聞き、敬語の誤り、言い残したことはないかどうかのチェックなどを全員で確認していく。最後に筆者が作成した例文を与えて確認。この際、使用される敬語に特に注意させた。

②「練習アクティビティー」

筆者から渡された例文をペアーで役割を交代しながら読んでみる。次に、例文を基

例1　学生による手紙の一例（縦書きを横書きに直したもの）



拝啓　突然お手紙を差し上げます失礼をお許しください。私どもは、ハバロフスク国立教育大学の四年生の者でございます。

本日は、折り入ってお願いがございます。実は、今、当校では「ハバロフスクで働く日本人」というテーマで勉強をしております。つきましては、ぜひ岡田様にそういったテーマで御意見をうかがいたいと思っております。食事や仕事、また日本人とロシア人の労働観の違いなどの問題についていろいろ教えていただけないでしょうか。

十二月中旬にインタビューを予定しております。　改めて本状がお手元に届きます頃を見はからってご連絡を申し上げ、御意向をうかがう所存です。

ご多忙の折りから恐縮に存じますが、どうかよろしくおはからいくださいますようにお願い申し上げます。

右、取り急ぎインタビューのお願いまで。

敬具

一九九四年十二月四日

ポストニコフ
ブルロワ
チナナリョーワ

岡田一[illegible]様

本にしながら一人がインタビューを受ける側、一人がする側になり、配布されたお互いの異なる予定表を見ながら日時を調節し、会う場所を決めるロールプレーを行う。その後、いくつかのグループに代表で会話をさせた（例２参照）。

③「実際使用アクティビティー」

「練習アクティビティー」の２、３日後の授業時間外に学校の職員室から実際にインタビュー依頼の電話をかけさせる。その場には電話をかける代表者の他に筆者、他のメンバーなども立ち会う。電話の後、筆者は簡単に学生にコメント程度の評価をした。

「インタビュー段階」

①「解釈アクティビティー」と②「練習アクティビティー」

ここでは二つのアクティビティーが交互に行われた。

まず、インタビュー場面において予想される全体的な流れを示したプリントを渡し、次にそれぞれの段階で必要と思われる表現等の確認をしていった。具体的には、挨拶、自己紹介の仕方、考え・意見・感想の聞き方、相槌、言い淀みの種類や機能、話題の変え方、コミュニケーション問題が起きた時の処理の仕方（聞き返し、言い換えを求める、相手の言葉を自分の言葉で言い換える、相手の言葉を確認する）、話を終えるときの方法などを以下の教材を利用しながら説明、練習を繰り返していった。

例２　電話での参考会話例

（本人の場合）
A：もしもし、（会社or本人の名前）です
が。
B：あ、もしもし、私、ハバロフスク教育大
学の（自分の名前）と申しますが、～さんを
お願いできますでしょうか。
A：はい、私ですが。
B：あ、始めまして。
A：始めまして。

（本人でない場合）
C：もしもし、（会社or本人の名前）ですが。
B：あ、もしもし、私、ハバロフスク教育大学の
（自分の名前）と申しますが、～さんをお願いで
きますでしょうか。
C：はい、少々お待ちください。
...
A：もしもし、お電話代わりました。A　です。
B：あ、私、ハバロフスク教育大学の（自分の名
前）と申します。始めまして。
A：始めまして。

B：あのう、先日私どもが差し上げましたお手紙はお手元に届きましたでしょうか。
A：ああ、はい、いただいています。何かインタビューをなさりたいとか...。
B：はい、実は今度、日本語のクラスで、「ハバロフスクで働く日本人」というテー
マで日本の方にインタビューをして、その結果を発表することになったんです。それ
で、～さんにお仕事の内容や、日本人とロシア人の労働感の違いなどについて三十分
程質問をさせていただきたいと思いまして...。
A：ああ、いいですよ。喜んでいたしましょう。
B：あ、どうもありがとうございます。あのう、それで、インタビューの日時と場所
なんですが、私どもは（　　）の（　　）時ごろ、そちらにうかがおうかと考えてい
るのですが、Aさんのご都合はいかがでしょうか。
A：そうですねえ、その時間はちょっと...。（　　）の（　　）時ごろはどうです
か。
B：ええと、ちょっとお待ちください。...あ、はい、結構です。それでは（　　）の
（　　）時に、同じ教育大学のDとEというものと三人で参りますので、よろしくお願
いいたします。
A：はい、わかりました。それじゃあ、お待ちしています。
B：はい、それではどうもありがとうございました。失礼いたします。
A：さようなら。

＜教材：①「インタビューで学ぶ日本語」（凡人社）②「ロールプレーで学ぶ会話1＆2」（凡人社）③「待遇表現」（ジャパンタイムズ）＞（例３参照）

「練習アクティビティー」では主にペアーで会話文を読ませたり、またドリル、インタビューのロールプレーなどを行った。

また、社会文化的な側面、例えばお辞儀の仕方や勧められるまで椅子に座らない、足を組まないなどの確認もした。

③「実際使用アクティビティー」

各グループが決められた日時、場所に出かけて行ってインタビューをした。また、インタビューの内容は録音された。

「発表段階」

①「解釈アクティビティー」と②「練習アクティビティー」

対象となった学生達には新学期開始時から、発表や司会を行う活動を定期的に導入

していたので、今回は特に改めて何も行わなかった。しかし、確認のために、司会の始めの言葉、インタビューを受けた人の紹介、発表者の紹介、質問に対する答え方、感想の述べ方などをプリントして渡し、目を通させた。

③「実際使用アクティビティー」

グループごとにインタビューの結果を報告した。発表時間は各グループ２５分程度で、３人のうち１人が司会者と発表者を兼ね、残りの学生は発表だけをした。また、他のグループの発表に対して必ず一つは質問をするように求めた。この発表内容は録音された。

効果

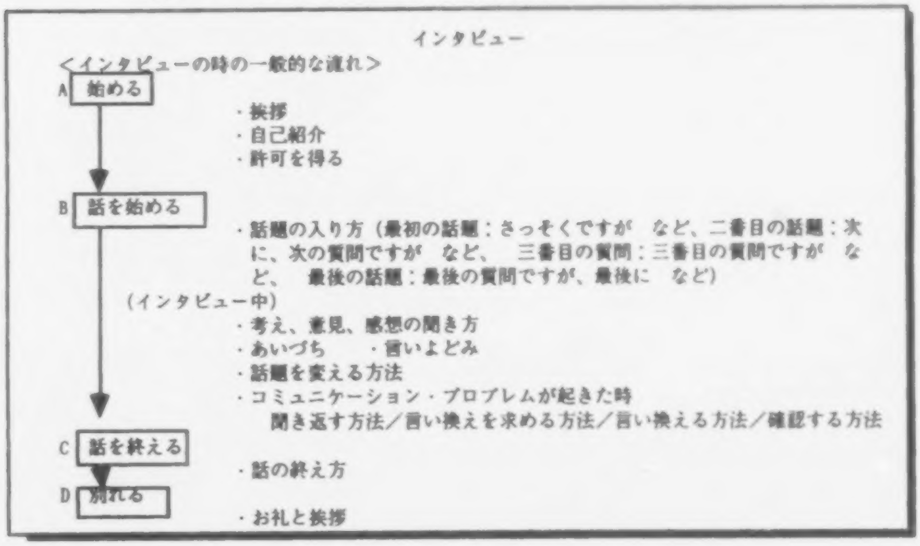
ここで挙げられる効果とは基本的にインタビューを行ったことによる学習効果のことで、インタビューとその発表内容、それにコース終了後の各学生に対するフォローアップインタビューの分析結果からわかったことである。

言語学的効果

この節では、インタビュー時、あるいはインタビューをきっかけにして、習得されたと思われる語彙、表現の例を挙げる。

習得されたかどうかの判断は、インタビュー後の筆者との自然な会話の中で適切にそれらを使用していた場合、「発表段階」でそれらの意味の説明を行ったり、正しい文脈の中で使用していた場合、また学生とのフォローアップインタビュー中に筆者にそれらの意味を適切に説明できた場合のみを考慮に入れた。

例３　インタビューの流れとそれぞれの段階の学習事項


インタビュー
<インタビューの時の一般的な流れ>
A 始める
・挨拶
・自己紹介
・許可を得る
B 話を始める
・話題の入り方（最初の話題：さっそくですが　など、二番目の話題：次に、次の質問ですが　など、　三番目の質問：三番目の質問ですが　など、　最後の話題：最後の質問ですが、最後に　など）
（インタビュー中）
・考え、意見、感想の聞き方
・あいづち　　・言いよどみ
・話題を変える方法
・コミュニケーション・プロブレムが起きた時
聞き返す方法／言い換えを求める方法／言い換える方法／確認する方法
C 話を終える
・話の終え方
D 別れる
・お礼と挨拶

また、習得されたと思われる語彙や表現が実際に学生にとって新しいものであったかどうかは、フォローアップインタビューの際の学生の証言によった。

①インタビュー中、会話の文脈から推測し、新しい表現の習得のきっかけを得た例。

[1]「郷に入っては郷に従え」

学生Aは、日本人がインタビュー中に使用したこの表現の意味を会話中に推測し、理解したと述べている。学生Aはインタビュー後、実際に筆者との会話の中でこの表現を数回適切に使用している。

[2]「日本人のお客はサービスにうるさい。」

この表現は学生Bにとって初めての表現だったが、文脈からその意味を推測し、その後発表までの準備の過程で完全に理解したと述べている。また、学生Bは発表の際にこの表現を適切に使用している。

②インタビュー中の日本人側から学生側への語彙知識の確認、あるいは学生側から日本人側への説明要求が日本人による語彙の説明を導き、その結果それらの語彙が習得されたと思われる例。学生はこれらの語彙の意味をフォローアップインタビューの際、筆者に適切に説明している。

語彙知識の確認の例：「保険、わかりますか？」、「営利活動ってわかる？」
説明要求の例　　　：「責任感ってなんですか？」

③インタビュー中に「予備段階」で学習した語彙が現われ、習得が強化された例。以下の語彙も"と同様、筆者は学生とのフォローアップインタビューの中でそれぞれの語彙に対する学生の理解を確認している。

「残業」、「週休二日制」、「働き蜂」

④インタビュー中の意味不明の単語などを後に辞書で調べる過程で習得した例。

「楽天家」（この語を習得した学生B は、後に筆者との日常会話でも実際に使っている）、「利息」（この語を習得した学生C は発表時に他の学生からの質問に対して日本語で説明をした)

社会言語学的、社会文化的効果

①日本語に存在する多様性を知った。

青森県出身の会社員にインタビューを行った学生達は、その日本人のイントネーションが今までに接触した日本人と異なることに気付いたとフォローアップインタビューの際に語っている。また、この日本人は相槌の回数もかなり多く、同グループの一人の学生はこの相槌の個人的特徴にも言及している。

②日本人の視点から自分達の国を見ることができた。

以下の文は日本人によって発話されたものである。

「…ロシアの銀行はそういった意味で（日本の銀行と比べると）まだまだ、真似事…」

「ロシアでは、お店で袋がもらえませんね。」

「ロシアの旦那さんは奥さんの手伝いをしないんですか！」

上記のものは学生とのフォローアップインタビュー、あるいは後の筆者との会話の中で学生が特に印象を受けたと筆者によって判断されたものである。いずれも学生達にとって、日本人がロシアに関して考えていることや驚いたことを知ることができ、同時に、それらを通して逆に日本の様子や日本人の自国に対する考え方を学ぶことができた。

教育上の効果

Strevens (1987, p. 171)は、教室外での母国語話者との実際のインターアクションを授業に取り入れることによって、学生にとって授業が興味深いものとなる、学生が会話者としての役割を自覚する、学習意欲が向上するなどの効果がもたらされるとしている。また、Dickinson (1995, p. 174)によれば、学生に自らの学習に対する責任を与えることは、動機付けを促進し、学習を成功に導く条件の一つであるとしている。

今回のこのプロジェクトでも、学生達は教室外での実際のインターアクション場面を与えられ、また同時にこのインターアクション場面では教師の監督下を離れ、自ら言語使用の検証、修正を行っていく責任をも与えられた。このことは学生のより積極的、自発的な学習計画全体への参加態度を生んだようだ。また、マンネリ化した教室場面から離れることも学習意欲向上の一原因となったと思われる。

問題点

ここでの問題点とは5章同様、「インタビュー段階」を中心に筆者によって認識された問題点であり、学生へのフォローアップインタビュー、及びインタビューを受けた日本人側へのフォローアップインタビューを通して明らかになったものである。

言語学的問題点

インタビューの結果、間違った日本語を記憶してしまった学生がいた。

「るっか」、「えんたい、たいえん」

学生Eは発表時に意味不明の上記のような日本語を使用した。彼女はこれらを経済の専門用語であると理解していたが、フォローアップインタビューの際、一緒にテープを聞き直してみると、いずれも彼女が理解可能な語彙だった。つまり、「るっか」とは「物価」のことであり、「えんたい、たいえん」とは「円対ルーブル」の意であった。この場合、相手の日本人の方言（青森弁）や話し方が原因になったと思われる。

インタビューの技術上の問題点

①仕事に関する質問、生活に関する質問が交互に現われ、内容ごとに質問の順序が組み立てられておらず、インタビュー全体にまとまりが見られないグループがあった。この場合、日本人は矢継早にいろいろ質問を投げかけられているといった印象を持ったようだ。

②それぞれ分野の違う日本人に仕事に関する内容を聞く場合、当然その分野に関係する専門用語が出てくるはずである。したがって、それに対する各グループごとの事前の準備、または相手の答えの中に含まれる語彙あるいは内容の難度を下げるための質問作成段階でのコントロールが必要だった。あるグループは日本人の銀行家に「ハバロフスクの経済状態についてどう思われますか。」という質問をし、かなり難しい説明を受けた。そのため学生達にとってそのインタビューがほとんど内容のわからない、つまらないものとなってしまった。しかし、彼等は相槌だけは打ち続けたため、その銀行家はわかっているものと思い込み、ほとんど調整を加えずに話し続けた。この場合、説明要求などのストラテジーの使用も必要だったであろう。

③８グループ中３グループが録音に失敗した。録音は、後に評価をするための重要な材料となるものなので、録音機の状態を学生に前もってよく確認させておく必要があった。

その他

前述したようにインタビューを受けた日本人のうち４人に簡単なフォローアップインタビューを行ったが、すべての日本人についてさらに詳しく調査する必要があった。日本人側が不快に思ったり疲労感を覚えたりした状況を分析することは、次回の指導のポイントの有効な指標になりうる。また、これはインタビューだけではなく、手紙や電話の段階でも、できるだけ行なわれる必要があった。

おわりに

今回のプロジェクトは言語習得の促進を図り、学習に対する動機付けを強化し、さらに日本語や日本人の考え方の多様性を認識させるなどの点で一応の効果を達成したものと思われる。しかし、その分析対象が「インタビュー段階」の主に「実際使用アクティビティー」を中心としており、今後は各段階における３つのアクティビティーの流れが有機的にもたらす効果、また全段階の統合の結果による効果を詳しく検証する方法を考えていかなければならない。同時に、学生に対する評価に関しても「発表段階」以外の段階をどのように適切に評価していくかを検討していく必要がある。

また、自然なインターアクションの中で学生の日本語運用能力の向上を目指す場合、この様なプロジェクトを年に一度実施するだけではさほど効果は期待できない。プロジェクトは長期にわたるものである必要は必ずしもなく、短期的なものでも今回の様な「解釈アクティビティー」、「練習アクティビティー」、「実際使用アクティビティー」という流れに沿ったものを定期的にカリキュラムに組み込んでいく必要があろう。

加藤好崇はオーストラリア、モナシュ大学日本語学科マスターコース終了後、１９９４年６月より１９９５年６月まで国際交流基金、日本外交協会による派遣日本語講師としてロシア連邦ハバロフスク国立教育大学日本語科に在籍した。現在、東海大学留学生センターに勤務している。

BEST COPY AVAILABLE

参考文献

Dikinson, L. (1995). Autonomy and motivation: a literature review. *System, 23*(2), 165-174.

Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Blackwell.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Richards, J., Platt, J., & H. Weber. (1986). *Longman dictionary of applied linguistics*. London: Longman.

Strevens, P. (1987). Interaction outside the classroom: using the community. In W. Rivers (Ed.) *Interactive language teaching* (pp. 170-176). Cambridge: Cambridge University Press.

尾崎明人・ネウストプニー,J.V. (1986).「インターアクションのための日本語教育：イマーションプログラムの試み」.『日本語教育』59号, 41-52.

ネウストプニー,J.V. (1981).「外国人場面の研究と日本語教育」.『日本語教育』45号, 30-40.

ネウストプニー,J.V. (1995).『新しい日本語教育のために』. 大修館書店.

1996年1月8日受
1996年6月20日改定版受
1996年11月1日改定2版受

Research Forum

Japanese EFL Learners' Test-Type Related Interlanguage Variability

Akihiro Ito
Hiroshima University

The purpose of this article is two-fold: 1) to investigate the effects of differences in test-types on the accuracy rates in interlanguage performance of Japanese EFL learners, and 2) to examine the reliability and validity of a grammaticality judgment test. Three grammar tests of relative clauses with three different test-types were assigned to 41 Japanese high school students. The tests were constructed on the basis of amount of attention to linguistic form from reviewing recent SLA works on task variation theory. From the results of the investigation, it is argued that 1) the participants showed different accuracy rates in the three test-types according to the expected order, and 2) the grammaticality judgment test showed relatively high reliability (*rxx'*=0.792) and moderate validity. The article also discusses the pedagogical implications of the findings and future direction of research.

この論文の目的は二つある。一つは日本人英語学習者の中間言語パフォーマンスへのテスト・タイプの効果を調査することであり、もう一つは文法性判断テストの信頼性と妥当性を検証することである。3種類の関係詞節に関するテストを41人の日本人高校生に実施した。テストは最近の第二言語習得研究におけるタスク要因理論を参考にして、言語形式への注意の度合を変数として作成された。調査の結果、3種類のテストにおいて仮説通りの順番で正確さの度合が異なり、文法性判断テストは比較的高い信頼性（rxx'=0.792)と中程度の妥当性を示した。本稿はさらにこの研究の教育への示唆と将来の研究の方向についても論じている。

When we try to estimate language learners' competence by measuring performance on a certain grammatical item, we often find it quite difficult to decide which test-type should be used. Recent research on second language acquisition theory has suggested that the accuracy rates of interlanguage performance systematically vary according to the kind of test-type[1] (Kameyama, 1987; Ohba, 1994a, 1994b). This phenomenon has been mainly explored in

JALT Journal, Vol. 19, No. 1, May, 1997

an area of second language acquisition research called task variation theory (Ohba, 1987).

In the early to mid-1990s, researchers in the field of language testing claimed that research on language achievement tests had been neglected (Weir, 1993; Negishi, 1995) because most language testing professionals were more interested in measuring learners' general or overall proficiency.[2] Researchers tried to construct tests with potential for explaining learners' proficiency. As a consequence, little attention was being paid to test-type.[3] From the late 1970s through the 1980s, findings from SLA research implied that it was dangerous to measure learners' achievement level on a grammatical item or feature through use of a test containing only one test-type[4] (c.f. Kameyama, 1987; Nunan, 1992; Ohba, 1994a, 1994b). This was because of the reported systematic performance variability shown in foreign or second language learners' interlanguage according to test-type (Tarone, 1979, 1982, 1983, 1985, 1988; Sajjadi & Tahririan, 1992). Hence, the accuracy rate of the learners' interlanguage performance measurements can be affected by the test-type even if intended to measure a similar trait (Bachman, 1990; Ohba, 1994a, 1994b). This has served as the motivation for the present research. In this study relative clauses are the subject for three different test-types constructed to examine the effect of difference in test-type on Japanese EFL learners' interlanguage performance.

Test-type Classification

Recent research on test-types in SLA began with the dichotomous catergorization of test types into Natural Communication Tasks and Linguistic Manipulation Tasks (Dulay, Burt, & Krashen, 1982). A Natural Communication Task required learners to pay attention to content in order to use language for communication. A Linguistic Manipulation Task asked learners to pay attention to linguistic manipulation of form. This categorization was based on the *Monitor Model* (Krashen, 1977a, 1977b, 1978a, 1978b, 1981, 1982; Krashen & Terrell, 1983). The *Monitor Model* "predicts that the nature of second language performance errors will depend on whether monitoring is in operation" (Krashen, 1982, p. 152). Therefore, the Linguistic Manipulation Task, which encourages use of the monitor, measures to what degree learners have mastered grammatical rules and permits them to show higher accuracy. On the other hand, the Natural Communication Task, which does not permit monitor use, measures subconsciously-learned grammatical rules, which results in lower accuracy. In other words, Dulay, Burt, & Krashen (1982) suggest that test-types be classified rather like an on-off switch based on whether the test-type allows the monitor to be in operation or not.

Criticisms of test-type classification based on this dichotomy were quick to appear. Tarone (1983) first modified the dichotomy of test-types from the Labovian sociolinguistic perspective, arguing that there was no clear on-off point. She then proposed that the accuracy of interlanguage performance ranges on a continuum from vernacular to careful style based on the *Interlanguage Continuum Model*. Following this model, Hyltenstam (1983) proposed a more detailed categorization of experimental data, suggesting the following eight test-types:

> a) Elicited production, often with pictorial stimuli, such as the *Berko Test*, the *Bilingual Syntax Measure*, guided composition;
> b) Manipulation of given linguistic material such as sentence combining and sentence completion;
> c) Intuition or grammaticality judgment test;
> d) Introspection;
> e) The cloze procedure;
> f) Imitation;
> g) Dictation or partial dictation; and
> h) Translation. (p. 58)

Hyltenstam (1983), examining research on the relationship between interlanguage performance and selection of test-types, argued that it was impossible to divide test-types into two groups and reasonable to categorize them according to how much each required attention to linguistic form and content. According to this classification, as learners move toward translation, and pay greater attention to linguistic form, the resulting accuracy rate in the test is higher.[5]

Research on Interlanguage Variability

Most research on the effect of test-types on learners' interlanguage performance has not stemmed from research on testing but from other perspectives of SLA (Ohba, 1994a, 1994b). Ohba (1987, 1994a, 1994b) makes a rough division of what kind of grammatical items are chosen for the purpose of second language acquisition research, categorizing them into three types. The first, concerned with phonology, investigates the relationship between accuracy rate of a particular phoneme and the type of tasks provided to the participants (Dickerson, 1975; Beebe, 1980; Sato, 1985; Schmidt, 1987; Shizuka, 1993). The second examines the acquisition of morphology (Larsen-Freeman, 1975; Kameyama, 1987; Tomita, 1988; Inoi, 1991; Takamiyagi, 1991). The third discusses the acquisition of syntax (Bailey, Eisenstein & Madden, 1976; Schmidt, 1980; Hueber, 1985). The results from these studies have supported the idea that the accuracy rate of language learners' performance changes according to the *Interlanguage Continuum Model,* though some morphological features such as articles have been shown not to follow it.

Recently, Ohba (1994b) investigated the effect of different test types on the interlanguage performance of Japanese EFL learners. The noteworthy points in this study were:

1. the use of a larger number of participants (*N*=370) in order to generalize the findings to a statistical population;
2. the division of participants by proficiency level (higher, average, lower) from scores on the *STEP* placement test (Obunsha, 1987, 1989) to determine whether higher-level learners show differences in accuracy rates;
3. the use of three test-types a) Grammaticality Judgment, b) Sentence Combining, and c) Picture Description;
4. the selection of a complex syntactic structure, relative clauses, to examine the effect of test-types because accuracy rates for morphology such as articles are easily influenced by discourse (Long & Sato, 1984; Tarone, 1985; Tarone & Parrish, 1988; Ohba, 1994a, 1994b), making it difficult to determine if the difference affects interlanguage performance.

Though Ohba (1994a, 1994b) made a significant contribution to this field, there were, nevertheless, a few drawbacks to his research. First, since different sentences were used for each test, it cannot be concluded that the scores were affected only by test-type. Second, in the Picture Description test, learners were asked to produce subject-type relativised sentences for sentences containing relative pronouns because these are considered easiest for L1 and L2 learners of English (Schachter, 1974; Keenan & Comrie, 1977; Gass, 1980) and avoided other kinds of relative clauses. The other test-types, Grammaticality Judgment test and Sentence Combining test, consisted of a mixture of sentences with four locations of the head noun phrase to be relativised: subject, direct object, object of preposition, and possessive. Therefore, though the Picture Description test was classified as not requiring attention to linguistic form, the characteristics of relative clauses may have resulted in higher than expected accuracy rates. Third, though Ohba said the low time pressure might have activated the learners' monitor in the Picture Description test to explain the accuracy rate, Krashen (1985) said that time pressure may be unrelated to monitor activation. Therefore, it is difficult to view lack of time pressure as the reason for the increased accuracy rate.

In response to these concerns, the effects of test-types alone on learners' interlanguage performance need to be examined.

The Study

Purpose and Research Questions

The purpose of the present study was threefold: first, to replicate and expand Ohba's (1994a, 1994b) findings under more strictly controlled conditions; second, to examine the reliability coefficients of the three tests, and third, to determine what the grammatical judgment test employed by Ohba examines.

It was expected that the participants' interlanguage performance would vary according to Hyltenstam's (1983) theoretical framework.

Grammaticality judgment tests tend to be criticized (Ohba, 1994b), since some researchers (Ellis, 1991, 1994; Chaudron, 1983) are skeptical of their reliability. Moreover, some researchers question what grammaticality judgment tests measure in comparison to tests with different test-types. In spite of such criticism, grammaticality judgment tests are widely used because they are easy to construct and are believed to be a useful tool for eliciting learners' linguistic knowledge (Ohba, 1994b). To address these concerns, a grammaticality judgment test constructed on the basis of findings concerning relative clauses (Gass, 1980; Kawauchi, 1988) was used to examine the reliability. In order to determine what the grammaticality judgment test really examines, the correlations between the grammaticality judgment test and other tests, which target the same trait, but with different test-types, were calculated.

The research questions are:

1. Does the accuracy rate of the three tests follow the pattern of: Cloze > Grammaticality Judgment > Sentence Combining?
2. What does the Grammaticality Judgment measure? Is this test's reliability low?

The small sample ($N=41$) necessitated a conservative treatment of statistical analyses. Therefore, the alpha level for all statistical decisions was set at $a<0.01$.

Method

Subjects: The subjects in this study (N=41) were second year high school students (10th grade) enrolled in an English reading class at a high school attached to the Aichi University of Education. All were native speakers of Japanese. The average age was 16. All had completed at lease four years of formal English courses. The sample was thus homogeneous with regard to nationality, language background, educational level, and age. The group consisted of 21 males and 20 females. One male student was absent throughout this study. Though he later took the tests, his scores were not included.

In general, Japanese students are required to learn the usage of relative clauses: subject (SU), direct object (DO), indirect object (IO), and genitives (GEN), in junior high school, and relative clauses of preposition (OPrep) in high school. Therefore, it was concluded that the Ss had basic ability regarding use and comprehension of the English relative clause.

Materials: The following three 24-item relative clause tests were administered: 1) Cloze, 2) Grammaticality Judgment, and 3) Sentence Combining. These tests are a modification of the tests in Ohba (1994b) (see Appendix). To prevent use of only the easiest type of relative clause, subject type (SU), the same sentences were used for each test and the Picture Description test was replaced by the Cloze.[6]

In the process of constructing the three relative clause tests, careful attention was paid to the Noun Phrase Accessibility Hierarchy: SU (subject) > DO (direct object) > IO (indirect object) > OPrep (object of preposition) > GEN (possessive) > OComp (object of comparative particle) (Keenan & Comrie, 1977). The location of the head noun phrase is considered an influential component in the degree of difficulty associated with relative clauses in both L1 and L2 English acquisition (c.f. Schachter, 1974; Keenan & Comrie, 1977; Gass, 1980; Eckman, Bell, & Nelson, 1988; Akagawa, 1992; Sadighi, 1994; Aarts & Schils, 1995).

The tests: The 24 questions in the Sentence Combining test contained six pairs of sentences to be combined into sentences containing a relativised SU, six into sentences containing a relativised DO, six into sentences with a relativised OPrep, and six into sentences containing genitive cases. In the Cloze test, an appropriate relative pronoun, who, which, whose, whom, was required. In the Grammaticality Judgment test, a determination of each sentence's grammaticality, using either "O" or "X" as markers for correctness and incorrectness respectively, was required. For sentences judged incorrect, Ss were asked to make necessary corrections.Typical errors in relative clauses are universal. They are categorized as: 1) relative clause marker omission, 2) pronoun retention, 3) wrong selection of relative clause marker, and 4) adjacency (Gass, 1980; Kawauchi, 1988). There were 12 correct and 12 incorrect sentences.

Test administration: In response to the shortcomings of recent studies, the ordering of the three test papers was taken into consideration. Since in the three tests the same sentences appear, the possibility of Ss memorizing the orthography to gain a higher score in later tests was considered. In order to reduce the potential order effect, the 24 items in each test were divided into three groups. Eight items from each of the tests were combined to make a 24-item test. In the first session, Cloze items 1-8, Grammaticality Judgment items 9-16, and Sentence Combining items 17-24 appeared; in the second session, Grammaticality Judgment items 1-8, Sentence Combining items 9-16, and Cloze items 17-24; and in the third session, Sentence Combining items 1-8, Cloze items 9-16, and Grammaticality Judgment items 17-24. Ss were allowed 20 minutes to complete each test. There was a one week interval between testing sessions, assumed to be long enough for Ss to forget some of the orthography and decrease the negative order effect. Ss were not told the dates of the sessions. Each session was conducted at the beginning of English reading classes by the instructor and his assistant. Though Ss were not informed of the purpose of the tests, they were encouraged to answer as many questions as possible. It is noteworthy that the Ss showed a great deal of interest on all the tests.

Scoring procedure: The tests were scored by the author. Scoring was based on whether the Ss had used the appropriate pronouns following an established criterion (Celce-Murcia & Larsen-Freeman, 1983; Quirk, Greenbaum, Leech, & Svartvik, 1985; Ohba, 1994a, 1994b). Therefore, local errors such as spelling mistakes were ignored as long as the meaning was clear.

Reliability estimation based on internal consistency: The Spearman-Brown split-half method was used to estimate the tests' reliability coefficients. The split-half method can be used when each test item is regarded as independent and also can contribute to the total score independently (local independence). In all three tests, each item is clearly independent. Thus, the split-half method is a permissible estimating procedure. The author scored the odd- and even-numbered items separately and first examined the Pearson's product-moment correlation (*r*) (Ito, 1996; Brown, personal communication, February 22, 1996). Each value was then corrected for the reduction to half-test length using the Spearman-Brown prophecy formula ($rxx' = 2r/1+r$).

Results and Discussion

In this section, descriptive statistics of the tests are shown and the research questions addressed. Before discussing the results, however, two other aspects must be considered:

1. the effects of sample homogeneity on reliability and estimated correlations among the three tests, and
2. the small size of the sample ($N=41$).

Table 1: Reliability Coefficient, Mean, Maximum Score & Standard Deviation of Tests ($N=41$)

Tests	Reliability rxx'	Mean (M)	Max. Score	SD
CL	0.596	14.976	24	4.891
GJ	0.797	12.976	24	3.309
SC	0.693	10.732	24	6.490

CL = Cloze test; GJ = Grammaticality Judgment test; SC = Sentence Combining test.

Table 2: Analysis of Variance Summary Table ($N=41$)

Source	SS	df	MS	F-ratio	p
Subjects	2333.659	40	58.341		
Test-Type	498.260	2	249.124	24.215	<0.000
Error	823.073	80	10.288		
Total	3654.073	122			

Table 3: Multiple comparison test (Ryan's method) summary table ($N=41$)

Pair	r	Nominal Level	t	p level	Significance
CP-SC	3	0.003	6.920	0.000	s.
CP-GJ	2	0.007	2.823	0.005	s.
GJ-SC	2	0.007	4.097	0.000	s.

CL = Cloze test; GJ = Grammaticality Judgment test; SC = Sentence Combining test.

1. Does the accuracy rate of the three tests follow the pattern of: Cloze > Grammaticality Judgment > Sentence Combining?

Table 1 shows that in the relative clause tests the Cloze test had the highest mean ($M = 14.976$), the Grammaticality Judgment test a lower mean ($M = 12.976$), and the Sentence Combining test the lowest ($M = 10.732$). The accuracy rate seems to have changed according to the expected order. In order to determine statistical significance, one-way analysis of variance was performed. Table 2 shows an overall significant difference in the three test scores ($F(2,80) = 24.833, p < 0.000$). A multiple comparison test using Ryan's method was performed to determine where the difference lay. Table 3 shows it existed in each pair of the three tests at $p < 0.01$. Therefore, hypothesis 1 was confirmed.

2. What does the Grammaticality Judgment measure? Is this test's reliability low?

Table 1 shows the reliability coefficients of the three relative clause tests. Unexpectedly, the Grammaticality Judgment test showed the highest reliability among the three tests ($r = 0.798$), with reliability high enough for it to be regarded as a reliable testing device. The Cloze test ($r = 0.597$) and the Sentence Combining test ($r = 0.694$) showed moderate reliability. Ranked from highest to lowest in reliability coefficients, the reliability was: Grammaticality Judgment > Sentence Combining > Cloze.

Table 4: Correlations between each pair of three tests ($N = 41$)

Tests	r (exploratory rate %)	p
CP & GJ	0.626 (39.189)	<0.01
GJ & SC	0.696 (48.442)	<0.01
SC & CP	0.698 (48.720)	<0.01

CL = Cloze test; GJ = Grammaticality Judgment test; SC = Sentence Combining test.

Results indicate that a Grammaticality Judgment test can be relatively reliable. However, validity also needs to be investigated. Table 4 displays the correlation for each pair of three tests. The correlations measured show almost the same magnitudes: $0.60 < r < 0.70$. Results reveal that each pair of tests shared the same amount of trait or ability needed for producing or comprehending relative clauses (exploratory rate or coefficients of determination: $r^2 = 39.189$ to 48.720%). [The square of value of r (r^2) *100 indicates what percentage of similar traits or abilities each pair of tests share]. The Grammaticality Judgment test shows a relatively high reliability coefficient and moderate correlation with the other two test-types.

Though Grammaticality Judgment tests are still controversial with regard to reliability and validity, this research indicates they have far-reaching potential as reliable and valid elicitation tools. However, test designers must be aware of the universal error types or typical errors from learners' L1 transfers in order to construct appropriate tests. In addition, Table 1 reveals that the Grammaticality Judgment test shows the lowest standard deviation here, implying it has limited discriminative ability.

Conclusions and Pedagogical Implications

The results of this study indicate that, for Japanese EFL high school students: 1) the accuracy rate follows the expected pattern (Cloze > Grammaticality Judgment > Sentence Combining); and 2) unexpectedly, the Grammaticality Judgment showed fairly high reliability, with moderate correlations between the two other test-types. However, since its discriminative ability seems limited, it should be used with extreme care.

In a pedagogical sense, the results indicate that the manifestation of learners' interlanguage competence, their performance, varies according to test-type. As a consequence, teachers may well underestimate or overestimate learners' knowledge or ability to use a grammatical item if they rely on only one test-type. Moreover, in order to characterize the learners' actual abilities in the target language, it is necessary to employ a variety of test-types.

Limitation and Suggestions for Further Research

It should be acknowledged that one of the limitations of the present investigation is that it focused only on the Ss' performance. Thus, the results can be generalized only for Japanese students. However, in many studies conducted in the past, various language backgrounds, ages, and educational backgrounds were mixed. As a result, the findings have often been hard to interpret because they can only be generalized to the single situation in which the data was collected. In addition, the results of this study may be influenced by two internal and external factors:

1. the nature of reliability in measures in general, and
2. restrictions in the range of ability that was sampled in the investigation.

Generally, tests are not simply reliable and valid but they can be reliable and valid for specific types of students and specific ranges of ability (Brown, personal communication, February 22, 1996). In this regard this research should be replicated with a larger sample of participants from a much wider population.

The following three general research questions are posed in the hope that other researchers will pursue further investigations.

1. Does the proficiency level affect the magnitude of inter-language variability with regard to accuracy rates? If so, does the degree of the variability in the target language decrease as proficiency level increases (i.e. higher level learners < average level learners < lower level learners?).[7]
2. What really causes the difference in accuracy rates? The amount of attention to linguistic form by monitoring? Or the difference in cognitive processes and demands required of subjects?
3. How high is the construct validity of each test? That is, does each of the tests measure what it is constructed to measure?

Acknowledgments

The author sincerely thanks Norihisa Matsumoto, Shogo Miura, Yoshito Nakamura, Tadashi Nishida, Nobukazu Matsu-ura, Hiroshima University, and J.D. Brown, University of Hawaii, for their valuable critical comments on earlier drafts of this paper. Thanks also go to Hiromasa Ohba, Health Sciences University of Hokkaido; Kenji Ohtomo, University of Tsukuba; and Susan Gass, Michigan State University for their communications on this topic, as well as Toshifumi Takizawa, Takao Imai, Yuya Koga, Shizuya Tara, Hiroyuki Araki, Hidetoshi Inoue, and Fumiyo Yoshikawa, for their encouragement and cooperation, and two anonymous reviewers of the JALT Journal *for their valuable comments. I alone, however, am responsible for the analysis herein.*

Akihiro Ito, M.A., Hiroshima University, is currently a Ph.D. candidate in the Department of English Language Education, Hiroshima University. His research interests include language testing, evaluation, and second language acquisition research methodology.

Notes

1. Throughout this paper, "test-type" refers to any task type. Tasks for data elicitation are widely used in language testing conditions. Terms such as test format, test method, or method facet can refer to the concept of test-type used here since there has been a great amount of variability in classification.
2. Following Bachman's (1990) definition, "performance" is used as an indicator of a long-standing ability or competence, a person's knowledge of the language, which can be estimated indirectly by a test score and its valid interpretation (p. 33). On the basis of Bachman's definition, I take the position that competence is basically homogeneous and unitary, and that interlanguage variability manifested in different test-types is essentially a phenomenon of performance. The variability of performance can be observed by examination of change in accuracy rate or score on tests.
3. Some language testing and/or curriculum experts (Brindley, 1989) claim that the distinction between proficiency and achievement is not clear-cut. However, in my view, an achievement test is given to language learners at the end of a program to check if they have mastered the targeted items or skills. In this sense, an achievement test is, as Alderson, Clapham, and Wall (1995, pp. 11-12) argue, similar to a progress test. Proficiency tests are not based on language programs or classes. The main purpose of proficiency tests is to examine overall language ability.
4. For overviews of research on task variation theory, see Ohba (1987, 1994b).
5. Tarone's (1983) categorization of elicitation tools is comprised of tasks which are more differentiated for Hyltenstam (1983). Other test-types such as cloze tests or translation test-types in Hyltenstam's classification of experimental data can be located on a continuum with regard to accuracy rate, though some must be placed in the careful-style area.
6. Some might question if the cloze test-type in Hyltenstam (1983) allows higher accuracy than grammaticality judgment tests or sentence combining tests. DeKeyser (1990), however, argues that the fill-in-the-blank test-type, which is similar to the cloze, works as a more valid measure of monitored knowledge. He discusses the effectiveness of the fill-in-the-blank test-type for examining monitored knowledge compared to the other test-types. In this regard, my interpretation of Hyltenstam's theoretical framework is a combination of Tarone's (1983) interlangauge continuum model and DeKeyser's proposals.
7. Reexamination of the data collected to determine the effects of proficiency level on the magnitude of variability of accuracy rates in participants' interlanguage performance is currently underway.

References

Aarts, F., & Schils, E. (1995). Relative clauses, the accessibility hierarchy and the contrastive analysis hypothesis. *International Review of Applied Linguistics, 33*, 47-63.

Akagawa, Y. (1992). Avoidance of relative clause by Japanese high school students. *JACET Bulletin, 23*, 1-18.

Alderson, J.C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bailey, N., Eisenstein, M., & Madden, C. (1976). The development of wh-questions in Adult second language learners. In Faneslow, J., & Crymes, R. (Eds.), *On TESOL '76* (pp. 1-9). Washington, DC: TESOL.

Beebe, L.M. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning, 30*, 178-194.

Brindley, G. (1989). *Assessing achievement in the learner-centered curriculum*. Sydney: National Centre for English Language Teaching and Research.

Celce-Murcia, M., & Larsen-Freeman, D.E. (1983). *The grammar book: An ESL/EFL teacher's course*. Rowley, MA: Newbury House.

Chaudron, C. (1983). Research on metalinguistic judgments: A review of theory, methods, and results. *Language Learning, 33*, 343-377.

DeKeyser, R. (1990). Towards a valid measurement of monitored knowledge. *Language Testing, 7*, 147-157.

Dickerson, L.J. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly, 9*, 401-407.

Dulay, H., Burt, M., & Krashen, S.D. (1982). *Language two*. Oxford: Oxford University Press.

Eckman, F.R., Bell, L., & Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics, 9*, 1-20.

Ellis, R. (1991). Grammar judgments and second language acquisition. *Studies in Second Language Acquisition, 13*, 161-186.

Ellis, R. (1994). Data, theory, and applications in second language acquisition research. In Ellis, R. (Ed.), *The study of second language acquisition* (pp. 669-691). Oxford: Oxford University Press.

Gass, S. (1980). An investigation of systematic transfer in adult second language learners. In Scarcella, R.C., & Krashen, S.D. (Eds.), *Research in second language acquisition* (pp. 132-141). Rowley, MA: Newbury House.

Hueber, T. (1985). System and variability in interlanguage syntax. *Language Learning, 35*, 141-163.

Hyltenstam, K. (1983). Data type and second language variability. In Rongbon, H. (Ed.), *Psycholinguistics and foreign language learning* (pp. 57-74). Abo (Turku), Finland: Abo Akademi.

Inoi, S. (1991). Variation in interlanguage with special reference to articles and pronouns. *Annual Review of English Language Education in Japan, 2*, 1-10.

Ito, A. (1996). Testing English tests: A language proficiency perspective. *JALT Journal, 18*, 183-197.

Ito, A. [in press]. A study on the variability of test performance of Japanese EFL learners: A combination of two theoretical frameworks. *CELES Bulletin, 26*.

Kameyama, T. (1987). *Bunpo tesuto niokeru tesuto keishiki niyoru keitaiso seitoritsu no henka* (The change of accuracy rate in grammar tests by the difference of test-types). *CELES Bulletin, 17*, 248-252.

Kawauchi, C. (1988). Universal processing of relative clauses by adult learners of English. *JACET Bulletin, 19*, 19-36.

Keenan, E.L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*, 63-99.

Krashen, S.D. (1977a). The monitor model for adult second language performance. In Burt, M., Dulay, H. & Finocchiaro, M. (Eds.), *Viewpoints on English as a second language* (pp. 152-161). New York: Regents.

Krashen, S.D. (1977b). Some issues relating to the monitor model. In Brown, D.H., Yorio, C.A., & Crymes, R.H. (Eds.), *On TESOL '77* (pp. 144-158). Washington, DC: TESOL.

Krashen, S.D. (1978a). Individual variation in the use of the monitor. In Ritchie, W.C. (Ed.), *Second language acquisition research: Issues and implications* (pp. 175-183). New York: Academic Press.

Krashen, S.D. (1978b). The monitor model for second language acquisition. In Gingras, R.C. (Ed.). *Second language acquisition and foreign language teaching* (pp. 1-26). New York: Center for Applied Linguistics.

Krashen, S.D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.

Krashen, S.D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Krashen, S.D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S.D. & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon Press.

Larsen-Freeman, D.E. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*, 409-419.

Long, M.H., & Sato, C.J. (1984). Methodological issues in interlanguage studies: An interactionist perspective. In Davies, A., Criper, C., & Howatt, A.P.R. (Eds.), *Interlanguage* (pp. 253-279). Edinburgh: Edinburgh University Press.

Negishi, M. (1995). *Ri-dingu no tesuto to hyoka* (Reading test and evaluation). *Eigokyoiku (The English Teachers' Magazine), 1*, 29-31.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Ni kyu yosou mondaishu (Collections of Second Grade STEP Tests). (1987). Tokyo: Obunsha.

Ni kyu yosou mondaishu (Collections of Second Grade STEP Tests). (1989). Tokyo: Obunsha

Ohba, H. (1987). *Tasuku barieshun riron ni tsuite* (On the task variation theory). *CELES Bulletin, 17*, 43-49.

Ohba, H. (1994a). *Nihonjin eigo gakushusha no chukangengo kahensei: Tasuku keishiki no kanten kara* (Japanese EFL learners' interlanguage variability: With reference to task-type). *CELES Bulletin, 24*, 187-192.

Ohba, H. (1994b). *Tesuto keishiki no chigai niyoru eigo gakushusha no pafomansu no kahensei* (Task-related variability in interlanguage by Japanese EFL learners). *STEP Bulletin, 6*, 34-48.

Powers, D.E. (1982). Selecting samples for testing the hypothesis of divisible versus unitary competence in language proficiency. *Language Learning, 32*, 331-335.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Sadighi, F. (1994). The acquisition of English relative clauses by Chinese, Japanese, and Korean adult native speakers. *International Review of Applied Linguistics, 32*, 141-153.

Sajjadi, S. & Tahririan, M.H. (1992). Task variability and interlanguage use. *International Review of Applied Linguistics, 30*, 35-44.

Sato, C.J. (1985). Task variation in interlanguage phonology. In Gass, S., & Madden, C.G. (Eds.), *Input on second language acquisition* (pp. 181-196). Rowley, MA: Newbury House.

Schachter, J. (1974). An error in error analysis. *Language Learning, 24*, 205-214.

Schmidt, M. (1980). Coordinate structures and language universals in interlanguage. *Language Learning, 30*, 397-416.

Schmidt, R.W. (1987). Sociolinguistic variation and language transfer in phonology. In Ioup, G., & Weinberger, S.H. (Eds.). *Interlanguage phonology* (pp. 365-377). Rowley, MA: Newbury House.

Shizuka, T. (1993). Task variation and accuracy predictors in interlanguage phonology production. *Bulletin of Kanto-Koshin-etsu English Language Education Society, 7*, 63-77.

Takamiyagi, T. (1991). A study of task related variation in interlanguage by Japanese university students in an instruction only environment. *Joetsu University of Education: Kenkyu Ronshu (Bulletin of Language Studies), 6*, 47-64.

Tarone, E. (1979). Interlanguage as chameleon. *Language Learning, 29*, 181-191.

Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning, 32*, 69-84.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*, 142-164.

Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning, 35*, 373-403.

Tarone, E. (1988). *Variation in interlanguage*. London: Edward Arnold Publishers.

Tarone, E. & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning, 38*, 21-44.

Tomita, Y. (1988). *Nihonjin kokosei no chukan gengo nitsuite no ichikousatsu: Keitaiso shutoku junjo kenkyu* (A study on the high school students' interlanguage: With reference to the acquisition order of morphologies). *CELES Bulletin, 18*, 208-213.

Weir, C.J. (1993). *Understanding and developing language tests*. London: Prentice-Hall International.

(Received Jan. 3, 1996; revised Oct. 7, 1996)

Appendix: Test sentences

Cloze

1. The policeman has caught the girl ( ) stole the car.
2. The author ( ) books I haven't read yet is well known.
3. The magazine from ( ) I got the information is Newsweek.
4. He still had the pen ( ) I had given to him.
5. I came across some students ( ) names I couldn't remember.
6. She paid the man from ( ) she had borrowed the money.
7. The wine ( ) you brought to our party was excellent.
8. You will receive the kind letter ( ) your mother wrote.
9. The book to ( ) I referred can be obtained from the library.
10. The girl ( ) is sitting at the reception desk is pretty.
11. The city from ( ) that boy came is far from here.
12. I met a friend ( ) mother was a famous designer.
13. The building ( ) stands near the lake is our hotel.
14. The man ( ) feet were very large has just bought new shoes.
15. I need someone ( ) can help me clean the house.
16. The woman ( ) John will marry next month is Japanese.
17. The man ( ) was injured in the accident is in the hospital.
18. I saw a man ( ) bag was the same as mine.
19. I have just found the key ( ) I lost yesterday.
20. The book ( ) I bought the other day is interesting.
21. The woman with ( ) he was talking was Mrs. Miller.
22. The police interviewed the lady from ( ) the diamonds had been stole.
23. The boy ( ) essay I corrected has entered Hokkaido University.
24. I know the children ( ) are playing in the yard.

Grammaticality Judgment

1. ( ) The policeman has caught the girl stole the car.
2. ( ) The author whose books I haven't read yet is well known.
3. ( ) The magazine which I got the information from it is Newsweek.
4. ( ) He still had the pen which I had given to him.
5. ( ) I came across some students names I couldn't remember.
6. ( ) She paid the man from whom she had borrowed the money.
7. ( ) The wine was excellent which you brought to our party.
8. ( ) You will receive the kind letter whose your mother wrote.
9. ( ) The book to which I referred can be obtained from the library.
10. ( ) The girl who is sitting at the reception desk is pretty.
11. ( ) The city is far from here which that boy came from.
12. ( ) I met a friend which mother was famous designer.
13. ( ) The building which stands near the lake is our hotel.
14. ( ) The man has just bought shoes whose feet were very large.

15. ( ) I need someone who can help me clean the house.
16. ( ) The woman whom John will marry next month is Japanese.
17. ( ) The man is in the hospital who was injured in the accident.
18. ( ) I saw a man whose bag was the same as mine.
19. ( ) I have just found the key which I lost yesterday.
20. ( ) The book which I bought it the other day is interesting.
21. ( ) The woman with whom he was talking was Mrs. Miller.
22. ( ) The police interviewed the lady from which the diamonds has been stolen.
23. ( ) The boy whose essay I corrected has entered Hokkaido University.
24. ( ) I know the children which are playing in the yard.

Sentence Combining

1. The policeman has caught the girl. She stole the car.
2. The author is well known, I haven't read his books yet.
3. The magazine is Newsweek. I got the information from it.
4. He still had the pen. I had given it to him.
5. I came across some students. I couldn't remember their names.
6. She paid the man. She had borrowed the money from him.
7. The wine was excellent. You brought it to our party.
8. You will receive the kind letter. Your mother wrote it.
9. The book can be obtained from the library. I referred to it.
10. The girl is pretty. She is sitting at the reception desk.
11. The city is far from here. The boy came from it.
12. I met a friend. His mother was a famous designer.
13. The building is our hotel. It stands near the lake.
14. The man has just bought new shoes. His feet were very large.
15. I need someone. Someone can help me clean the house.
16. The woman is Japanese. John will marry her next month.
17. The man is in hospital. He was injured in the accident.
18. I saw a man. His bag was the same as mine.
19. I have just found the key. I lost it yesterday.
20. The book is interesting. I bought it the other day.
21. The woman was Mrs. Miller. He was talking with her.
22. The police interviewed the lady. The diamonds had been stolen from her.
23. The boy had entered Hokkaido University. I corrected his essay.
24. I know the children. They were playing in the yard.

Codeswitching in EFL Learner Discourse

Ethel Ogane
Tokyo YMCA

L1 use by L2 learners has been couched in negative terms like "resort to" and "fall back on." However, L1 use can be looked at in a more positive way. The alternative use of the learners L1 and their approximation of the TL may be termed codeswitching (CS). This study looks at CS in five female Japanese EFL learners. Analysis of spoken data and insights from subjects' verbal reports suggest that CS is discourse related. CS helps learners manage and smooth the flow of conversation, and allows them to express their dual identities of student and individual in the classroom.

第二言語学習者による第一言語の使用は「第一言語に頼る」という表現にも見られるように否定的に捉えられてきた。しかしながら第一言語の使用はもっと肯定的に捉えることも可能である。第一言語と目標言語の近似値である中間言語との交替はコード切り替えと考えることもできる。本研究は、5人の日本人女性英語学習者によるコード切り替えを検討した。被験者の会話データと内省報告を分析したところ、コード切り替えは談話に関連していることが示唆された。コード切り替えは学習者が会話の流れを管理し円滑にする助けとなり、教室における学生と個人という二重のアイデンティティを表現する手段となっている。

Teachers may oppose the English Only movement on a sociopolitical level but accept English-only classroom management practices as common and natural (Auerbach, 1993). L1 use is couched in negative terms like "resort to" and "fall back on." In the literature on communication strategies, L1 use is labeled as a "compensatory strategy," problem solving in nature, used to overcome "insufficient linguistic resources" (Faerch & Kasper, 1983, p. 46). Tarone, Cohen and Dumas (1983) label it an "avoidance strategy" (p. 11). This attitude toward L1 use, as evidenced by the words "compensatory" and "avoidance," seems to focus negatively on learner interlanguage as linguistically inferior to the target language. Kasper and Blum-Kulka (1993) suggest that "instead of viewing interlanguage pragmatic knowledge and behavior as deficient in terms of native norms, we need to consider its functionality and inner justification" (p. 160).

JALT Journal, Vol. 19, No. 1, May, 1997

L1 use by learners may be described as codeswitching (CS) behavior. That is learners switching back and forth between their L1 and their approximation of the L2. Legenhausen (1991) suggests that CS in learners be compared and contrasted with that of more proficient bilingual speakers. Codeswitching, "the alternative use by bilinguals of two or more languages in the same conversation" (Milroy & Muysken, 1995, p. 7), was thought to be abnormal, random behavior and indicative of a lack of language proficiency. The last 25 years of research on CS has contributed to its "rehabilitation" (Gardner-Chloros, 1991, p. 59). This research has demonstrated that language alternation is not arbitrary behavior but a type of "skilled performance" (Myers-Scotton, 1993c, p. 47), an important and "extremely common characteristic of bilingual speech" (Grosjean, 1982, p. 146).

Types of CS

The types of switching identified in the literature include: *tag*, *intersentential*, and *intrasentential* (Appel & Muysken, 1987; Poplack, 1980; Romaine, 1995). Tag-switching is the placing of a tag, an exclamation, a formulaic expression, or a discourse particle from one language into an utterance which is, except for that item, entirely in another language. Tag-switching has been referred to as emblematic switching because the switched item serves to identify an otherwise monolingual utterance as bilingual in character. Examples[1] of tag-switches are:

1) *Vendia arroz* (He sold rice) 'n shit. (Poplack, 1980, p. 589)

2) I guess it's good *yo ne* (right)? (Nishimura, 1995, p. 169)

An intersentential switch occurs at clause or sentence boundaries in one of the languages or between speaker turns (Romaine, 1995), as in:

3) Sometimes I'll start a sentence in English *y termino en espanol* (and finish it in Spanish). (Poplack, 1980, p. 594)

Intrasentential switching, sometimes called code mixing, involves a switch within a clause or sentence and sometimes even within word boundaries. Examples are:

4) Why make Carol *sentarse atras pa'que* (sit in the back so) everybody has to move *pa'que se salga* (for her to get out)? (Poplack, 1980, p. 589)

5) And it's hard, 'cause me- *nanka, moo, bon o nanka yomu to,* cover-to-cover *yomanakattara*, if I stop *dokka de,* I forget the story. (And it's hard, because a person like me, when I read a

> book, unless I read it cover-to-cover, if I stop at some point, I forget the story.) (Nishimura, 1995, p. 178)

CS Research

The approach to CS has been interdisciplinary (Appel & Muysken, 1987; Milroy & Muysken, 1995; Tabouret-Keller, 1995). Researchers have attempted to locate and explain constraints on CS in terms of: a) how the bilingual mind works (psycholinguistic); b) the formal properties of linguistic systems (structural); and c) the social, historical and interactive processes of individuals and groups in language contact situations (sociolinguistic). Psycholinguistic aspects are not reviewed here.[2]

Structural aspects: Early studies tried to explain why particular points were chosen for switches. The free-morpheme constraint, proposed by Sankoff and Poplack (as cited in Romaine, 1995) predicted that "a switch may not occur between a bound morpheme and a lexical form unless the lexical form has been phonologically integrated into the language of the morpheme" (Romaine, 1995, p. 126). The equivalence constraint (Lipski, 1978; Poplack, 1980) held that a switch may occur where the juxtaposition of forms from two codes does not violate a syntactic rule of either code.

Recent work on CS has taken different approaches. The Matrix Language-Frame (MLF) (Myers-Scotton, 1993a) claimed production-based switching with constraints occurs not at the surface phrase structure but at an abstract level, the "mental lexicon" of the bilingual (p. 485). Romaine (1995), suggesting that "a mixed code has its own rules and constraints," said a switched item may not be "predictable from the individual constituent structure rules of the two systems in contact" (p. 160). Researchers also asked whether constraints were language-specific or universal (Romaine, 1995).

The study of linguistic constraints involves differentiating between borrowing and codeswitching. Grosjean (1982) defined borrowing as "a word or short expression that is adapted phonologically and morphologically to the language being spoken" (p. 308). However, problems of overlap in the actual diagnosis of whether an individual case is a switch or loan occurred, as when a speaker pronounces all words, borrowed or otherwise, in the same accent. Poplack (1988) maintained there is a dichotomy which could be worked out methodologically. Gardner-Chloros (1995) argued that CS is not "separable, either ideologically or in practice from borrowing, interference or pidginisation" (p. 86). Myers-Scotton (1993b) urged that codeswitching and borrowing be seen as part of a continuum.

Sociolinguistic aspects: CS was thought to be "part of the performance of the imperfect bilingual, motivated by inability to carry on a conversation" (Myers-Scotton, 1993c, p. 47) and not a serious research topic. A study of Norwegian dialects (Blom & Gumperz, 1972) presented CS as a legitimate topic of research and stimulated research on CS between languages (Myers-Scotton, 1993c). Two general kinds of CS were posited (Blom & Gumperz, 1972; Gumperz, 1982). In *situational* switching, speakers switch when they perceive or when there is a change in setting, topic of conversation, or participants. In *conversational* or *metaphorical* switching, speakers create meaning by switching. Gumperz (1982) suggested that CS is a discourse strategy used by bilinguals in much the same way monolinguals use style shifting and prosody. In addition, Gumperz (1982) distinguished between a "we code" and a "they code," usually the minority and majority languages respectively. Use of the "we code," implying intimacy and informality, is associated with in-group and personal activities, while the "they code," symbolizing authority and objectivity, marks more formal, less personal, out-group relations. Other CS functions included quotation, addressee specification, interjections, reiteration, and message qualification (Gumperz, 1982).

Other researchers have studied the functional aspects of CS. Poplack (1980) maintained that specific instances of switching in the smooth, rapid Spanish/English CS by Puerto Ricans in New York City cannot be assigned discourse functions. Rather, CS was the norm for this bilingual community. Poplack (1988) found a different pattern of CS among French/English bilinguals in the Ottawa-Hull community in Canada. These speakers drew attention to their CS by "repetition, hesitation, intonational highlighting" and "metalinguistic commentary" (p. 230). Language switching marked specific functions, including switching to provide an apt expression—*mot juste* switching (p. 228); to fill a lexical gap; to bracket or call attention to an English switch; and to explain, specify or translate.

The study of in-group speech repertoire of second generation Japanese-Canadian (Nisei) bilinguals (Nishimura, 1995) identified three speech varieties. In the basically Japanese variety, used with native Japanese speakers, CS serves to fill lexical gaps. In the basically English variety, used with each other, CS allows the speakers to express their shared ethnic identity. The mixed variety was used when the interlocutors included both native Japanese and Nisei. Nishimura (1995) categorized CS in the mixed variety into four groups: a) functions related to interactional (speaker/hearer) processes where Niseis tried to "reach out" (p. 167) to both types of listeners or to enhance rapport; b) discourse organization functions where CS marked the beginning or ending of a frame or introduced a topic; c) functions carrying stylistic effects where Niseis marked quotations with a switch in order to make their speech "more lively" (p. 177); and d) functionally neutral CS where the motivation was not clear.

Some researchers have attempted to move on from a descriptive approach listing CS functions to a prescriptive approach treating CS patterns within an explanatory framework or model. Scotton and Ury (1977), who looked at Swahili/Luyia/English CS, explained CS as an extension of a speaker in terms of the relationships between participants and subject and the social meaning of a language choice. By using a switch, the speaker may redefine the social arena, or set of norms, of an interaction. By continually switching, the speaker may avoid specifying the social meaning of the interaction. In her study of French/Alsatian CS, Gardner-Chloros (1991) showed how switching was used by colleagues in work situations to create solidarity in spite of gaps in Alsatian competence, especially for younger speakers, and the lack of appropriate technical terms in Alsatian. CS "is connected with individual factors which concern people's linguistic histories as well as their personalities" (Gardner-Chloros, 1991, p. 184).

CS Research in the Classroom

Early research into classroom CS concentrated on the "communicative functions of code-switching in teacher-led talk and on the frequency with which particular languages were employed to perform different functions" (Martin-Jones, 1995, p. 90-1). Later studies looked at the "sequential flow of classroom discourse" and the interactional work of teachers and students from a conversational analytic approach with an ethnographic grounding (Martin-Jones, 1995, p. 91). This included studies on CS in EFL and ESL classrooms. Martin-Jones (1995) cited two studies by Lin (1988, 1990) on CS in Anglo-Chinese secondary schools in Hong Kong. In her 1988 classroom study, Lin found that the teacher, a Cantonese/English bilingual, used CS "as a communicative resource to signal unspoken social meanings" (p. 84). Fotos (1994, 1995) looked at CS in EFL classrooms in Japan. In her analysis of Japanese/English CS by Japanese university EFL students, Fotos (1995) found that the learners used switching into Japanese for emphasis, repetition, or clarification, as well as to signal that a mistake had been made in English and a repair was to follow. These functions call attention to the speakers' English utterances.

Foreign and second language learners may be thought of as incipient bilinguals (Kasper, 1994) and EFL/ESL classrooms as bi/multilingual communities. This study analyzes CS in a group of Japanese adult EFL learners to further understand the use of CS. What functions might L1 (Japanese) use or CS serve for developing bilingual speakers?

The Study

Method

Subjects: Five female students, Maki, Remi, Emi, Fumi, and Kono (not their real names), comprising an elementary level "conversation" class at a small private language school voluntarily participated. All had studied English at the secondary level and two had had a year of post-secondary lessons. They had attended language classes for three to seven years. Their ages ranged from 45 to 69.

Data Collection: Data were collected from the subjects' class sessions, following Gardner-Chloros' (1991) suggestion that naturally occurring classsroom discourse rather than language elicited by experimental tasks is better suited for the study of CS. Data from retrospective oral interviews were also collected. Retrospective data were included for two reasons: first, this researcher hoped that the individual interviews would help raise the subjects' awareness of L1 use in the classroom and stimulate reflection, and second, that the verbal reports, "the learners' reports of their own intuitions and insight," would serve to "complement" (Cohen, 1987, p. 82) the classroom discourse data.

Class sessions: Fourteen two-hour class sessions were video and audio taped from May to October, 1994. This duration was selected to ensure that Ss would become accustomed to being recorded. Three sessions, from June and September, were transcribed. Only discourse involving the entire class was transcribed; pair work was not included. During pair work, the Ss inevitably spoke Japanese, but it was impossible to transcribe the separate pair conversations.

Transcription of the data is in standard English orthography. The Japanese switches, in italics, translated by the author[3] into idiomatic English, appear in parentheses. Significant contextual information appears in brackets.

Interviews: Each subject was individually interviewed twice. All interviews were audio-taped. Each interview took place directly after a lesson and lasted from one to two hours. Biographical data were collected in the first interview and retrospective data in the second.

During the first interview, Ss were asked about their English educational background and reasons for studying. Ss were told that they should feel free to use either Japanese or English and to request clarification whenever they did not understand. The researcher spoke mostly in English. During the interview, subjects were asked for comments on their use of Japanese in class.

During the second interview, the video tape of the lesson that had just taken place was shown. Subjects were asked to listen for instances when they used Japanese and to try to recall why they switched and what language they were thinking in when the switch occurred. Subjects were told that they could push the pause button whenever they wanted to make an observation or remembered something. I also tried to elicit comments from them on their CS.

Results and Analysis

Switches from one speaker's utterance to another's, and those within one speaker's utterance, were counted. Since the focus of this study is use of the L1, only switches from English to Japanese were considered.[4]

Table 1: Analysis of Switches to Japanese

Type of Switch	Frequency	%
Tags	39	8%
Interjections/Short-fixed Expressions	43	9%
wa/ga	3	1%
janakute	9	2%
Conjunctions	13	3%
ja	17	4%
Adjectives/Adverbs	18	4%
Nouns	45	10%
Fillers	48	10%
nan to iu	61	13%
Phrases/Sentences	164	35%
ni naru/ni suru	12	3%
Total	472	102%*

*Note: The Total equals more than 100% due to rounding.

Switches into the L1 designating place names and proper nouns were not included. Also excluded were loan-word status items like "bazaar" and "seminar" when they were used as attempts or approximations toward English. A phonological basis for determining whether a word was English or Japanese was not used since much of the subjects' interlanguage is highly accented.

As Table 1 shows, the learners frequently produced tag, interjection, noun, filler, and *nan to iu* (how do you say . . .) switches. Phrases and sentences were also a favored switch. Less fluent or L1 dominant bilinguals have been shown to prefer tag and single-item switches (McClure, 1981; Poplack, 1980). Fotos (1990), in her study of Japanese-English bilingual children's CS, found that single item and sentence switches were a "significant number of the total, 107 switches out of 153" (p. 84).

Tags: The most frequently used tag was *ne*. More than half of the tag switches were made by one learner, Fumi, who may have used *ne* as a mini-confirmation check (example 1) and as a repair acceptance marker (example 2). She also may have used *ne* to gain thinking time or keep her turn, saying she tried to use English as much as possible but felt that her English did not come out quickly enough.

1. Fumi: Matterhorn. But eh I can uh I can get to the, I could *ne*, I could get to the uh *eh* town . . .
2. Fumi: Healthness, healthy, mmm uh I I don't uh got a ill. I don't got after? uh after I
   Remi: After that.
   Fumi: After that, *ne*. After that I don't got the ill.

Subjects also used the tags *deshoo*, *kana*, and *ka*.

Interjections and short-fixed expressions: Switched interjections included *masaka*! (no kidding!), *naruhodo* (I see), and *e*? (what?); expressions included *gomen* (sorry) and *doomo arigatoo* (thank you). Other examples are:

3. T: I wore the earrings that I bought at the bazaar yesterday.
   Maki: *Maa!* (Oh!) [laughs] Many earrings you bought.
4. Fumi: One, two, three, four, five, six, seven, eight *ne*, plus eight.
   Kono: mmm
   Maki: *Sugoi!* (Wow!)

Maki and Remi reported in interviews that they were very aware of their Japanese classmates and sometimes used Japanese without thinking. Remi thought that she sometimes switched because she felt she had to respond in Japanese to her peers. Maki said she thought she used Japanese to be considerate of her classmates. The students appear to be using these switches to express personal feelings and to show solidarity with their peers. It is interesting to note, however, that in Example 3 Maki used an interjection switch in response to the teacher.

Wa *and* ga: There were three instances of *wa/ga* switches in the data. *Ga* is the subject marker and *wa* is the topic marker in Japanese.

5. Maki: Finish *ga*? (How about the finish?)
6. Emi: Finish *ga* October 14. (We finish on October 14.)
7. Maki: *Anoo* my lend *wa* very long time. Thank you very much. (I have borrowed this a very long time.)

Fotos (1994) also found *wa/ga* switches in her college student CS data and cites them as evidence against the free-morpheme constraint because the markers should be connected to Japanese topics and subjects. However, Fotos (1990) noted in her earlier children's CS study that these Japanese particles may not be bound "in the sense that bound morphemes exist in English or Spanish" (p. 88).

Nishimura (1989) found that second generation Japanese-American and Japanese-Canadian bilingual speakers, who all said they were better in English than in Japanese, stated a topic in Japanese then switched to a comment in English. The speakers' "thinking process is topic-comment" with the topic part realized in Japanese and the comment produced in English (p. 376). In Nishimura's (1995) Japanese-Canadian Nisei study, bilinguals speaking in the mixed variety marked topics with wa, using it to imply a contrast in order to "to single out something in a discourse" and "to indicate the reintroduction of a topic" (p. 174).

However, regarding the current study, the data are too limited to make a meaningful interpretation of the subjects' *wa* and *ga* switches. In addition, the grammaticality of example 7 is suspect.

Janakute *and* ja: In her study of Japanese college student CS, Fotos (1994) noted the students' use of *janakute* to signal that the preceding utterance was incorrect and would be repaired. She also noted use of *ja* to attract and focus attention on the utterance to follow. Similar instances were found.

8. T: So you had some time to think about your one minute speech. Okay, now who would like to be the first speaker?
SS: [laughs]
Maki: *Ja*, I'll try.
9. Fumi Uh there uh there is there there isn't there isn't uh penici penicillin *janakute* strep strep streptomycin mmm.

Conjunctions: Ss seemed to use conjunctions to help sort out their thoughts and get their messages out. As with tag switches, these may have been used to gain thinking time. Fumi said she wanted to speak English correctly so she was always thinking about whether or not she had made a mistake. Kono said that when something was difficult to express, Japanese seemed to just come out.

10. Fumi: Yes, Switzerland, uh there is there are uh four language in Switzerland, French, eh *eto* German
Kono: *Doitsu* (Germany). Germany.
Fumi: Germany
T: German
Kono: German
Fumi: Germany, *sorekara* (and), Germany, German, English and eh, Romansh.[5]
11. Kono: Yes, mm one time is Thursday night, mm *dakedo* (but) mm I have one time *dakedo* (but) mm I *nandakke na* (what was it) mm I hope is two time not not uh another schedule, is one time.

In Example 10, Fumi may be trying to gain time to think or work out the mistake. In Example 11 Kono may be using *dakedo* to get meaning across.

Adverbs and adjectives: Adverb and adjective switches appeared to function as ways to express personal feelings and show solidarity with classmates. In one instance Kono read a sentence and the other students had to guess who it was about.

12. Kono: started working thirty years a…go, go *ga okashii*. (go seems strange)
T: (laughs)
Emi: *Ja*, Kono [laughs]. (Well, I guess that's about you.)
Maki: *Tereteru* [laughs]. (Kono feels self-conscious.)

Maki expressed her feelings about Kono in Japanese. During the interview, Maki said that her first priority was not always to practice English but to communicate with friends.

Fillers: Fumi and Remi made most of the filler switches. During the interview, Fumi said she had not been aware of using *anoo* until asked about it. Remi used the word *nani* (what) as a filler and reported that it had the feeling of "uh" for her.

13. Fumi: K *anoo* Tomohiro Museum uh where where is *anoo*
Remi: Kusaki Dam? Near the Kusaki Dam.
14. Remi: How was your mother? Walk uh *nani* she took for a walk, uh took walk?

The filler switches may be functioning as a way to gain time and/or to show that the speaker needs help.

Nouns: Nouns were switched frequently. In a conversation about the summer water shortage, Maki explained how she tried to conserve water by giving plants the water she saved from washing rice (*kome*). She ended up using more water because she had to rinse out a rag (*zokin*).

15. Maki: Yes mmm but I I try the uh *kome* rice rice wash a *ko* uh rice. Uh, it's plant [laughs] but a roof roof of the uh wet I [laughs] had uh *zokin*.

For Maki in this instance, the Japanese word, *kome*, may have been more available than the English word, rice. While reporting on her performance in general, Maki said that she knew her ability was not very high and felt impatient while trying to communicate what she really wanted to talk about. She felt she had to explain herself more completely and add more even if it meant using an L1 word.

While talking about her trip to Switzerland, Fumi repeated "glacier lake" in Japanese. She may have thought that it was a difficult term for her classmates and switched to clarify the meaning.

16. Fumi: Switzerland very beautiful country, mountain and lake, eh glacier lake, eh *hyogaka no mizu*, very deep, deep green color.

Fumi said she sometimes uses a Japanese word because she hoped someone would give her an L2 equivalent, i.e. she fished for English words from the class.

Remi, while discussing a similar instance of noun switching (*minami juujisei* for the Southern Cross), said that she switched to Japanese to make sure that her classmates understood what she was talking about.

Nan to iu: Maki made more than half of the *nan to iu* switches. These switches seemed to have been made when the learners did not know the L2 word or could not come up with the item quickly enough.

17. Maki: . . . I tomorrow today morning this morning I take uh I my grandchild take a *hoikuen te nan to iu no*? (How do you say nursery school in English?)
18. Emi: Who is play?
Kono: Uh James uh mmm uh James *nandakke naa*? [laughs] Let's see, James what?)

The learners seem to be using these switches to gain time or request help with a word. Variations include *nandakke*, *te iu no*, and *te ieba ii no*.

Phrases and sentences: The Ss used numerous L1 phrase and sentence switches. They sometimes switched to Japanese to speak among themselves, effectively excluding this researcher from the conversation. During one lesson, the women seemed sleepy toward the end so we did some simple bending and stretching exercises to wake up. After the exercises this conversation took place:

19. Remi: *Kyoo wa sugoku tsukareru*. (I'm very tired today.)
SS: Uh, mmm [laughs]
T: Okay, now, okay.
Maki: *Kimari ga ii yoo na ki ga shite, koko de*. (I feel as if this would be a good time to end the lesson.)

Fotos (1994) found instances of CS where feelings are expressed in Japanese and factual information in English in her college student data. Similar instances were collected.

20. Remi: Raise your left hand and stretch.
T: Stretch. That feels good.
Remi: And put it down. Bend your *nani* um bend you *mae*. [laughs] *Zenbu wasuretteru*. (Bend your what um bend forward. I have forgotten how to say everything in English.)
21. Maki: Yes, maybe. [laughs]
SS: [fanning selves, whispering]

Maki: *Nugenai no konna no. Hazukashii.* (I can't take this off. I'm embarrassed.) [laughs] I have a plan.

In example 20 Remi gives commands in English and in 21 Maki answers in the TL, then both switch to the L1 to express feelings of helplessness and embarrassment.

Ni naru *and* ni suru: These intrasentential switches were very simple. They require little grammatical control and they do not violate the equivalence constraint.

22. Maki: Short break *ni suru*? (Shall we take a short break?)
23. Emi: Five weeks *ni naru*. (It will be five weeks.)

Discussion

The use of L1 switches such as tags, conjunctions, fillers, nouns, and *nan to iu* may help learners gain thinking time, smooth the conversation, get important points across, and signal for help. *Ja* and *janakute* switches may function to attract and focus attention on important L2 content and utterances. The subjects may be using interjection, adjective, adverb, noun, and sentence switches to express personal feelings and confirm their solidarity. The learners' CS might be labeled discourse-related switching (Martin-Jones, 1995), "speaker oriented" because it "serves as a resource for accomplishing different communicative acts at specific points within interactional sequences" (p. 99).

The motivations underlying the learners' language switching appear to be similar to some of the factors Gardner-Chloros (1991) gives for French/Alsatian bilingual CS in Strasbourg. Gardner-Chloros (1991) said CS is a kind of compromise

> between the exclusive use of one language and of the other, each with their respective cultural connotations; there are occasions, for example, when it seems too snobbish to speak French but too rustic to speak Alsatian and code-switching provides the solution. (p.184)

For EFL learners, it may seem too impersonal or difficult to speak English but too unlearner-like to speak Japanese. Thus, they codeswitch.

The motivations Nishimura (1995) outlined for Japanese/English Nisei CS might also be compared to those underlying these subjects' CS. Nishimura (1995) said the Niseis used their mixed variety when speaking to both native Japanese speakers and other Niseis to "reach out" to both types of speakers and "enhance rapport between speaker and the hearer" (p. 167-169). Japanese EFL learners may want or need to involve both the teacher and each other in communication, or they may be appealing to their dual identities of L1 speaker and L2 learner.

Legenhausen (1991) called CS a mode or register which learners feel

> to be the proper expression of their ambivalent psychological state, pitched between responsibility for their own learning process on the one hand, and their natural inclination to use their L1 in any communicative situation on the other. (p. 71)

The subjects here are in the classroom not just to learn English. They are individuals who appear quite aware of each other and who seem to want to get along. This may explain Fumi's switch:

24. T: [to Fumi] Comment, question or comment?
Fumi: Eh, okay, mmm eh, eh, *chotto saki ni, moo jikan ga mottainai.* (Please let someone else go ahead, I'm taking up too much time.)

As Burt (1990) found in her study of learner CS, "it is not always polite to be an extremely conscientious language learner" (p. 34).

It is important for teachers to keep in mind that foreign or second language learners are not becoming monolingual, they are becoming bilingual (Kasper, 1994). Our standards of competency and performance should include bi/multilingual models. There is also much to learn about the teaching and learning processes through the study of CS in the classroom. Martin-Jones (1995) argued that

> a conversational analytic approach to code switching in classroom discourse, grounded in ethnographic observation, can give … fine grained descriptions of the ways in which teachers and learners get things done bilingually in the classroom. (p. 103)

This study has been limited to learner CS in the classroom. Future studies should look at CS in the interactive discourse of teacher and students.

Notes

1. In the examples, the language other than English is in italics and an idiomatic translation follows in parentheses.
2. See Grosjean (1982) for an introduction to CS from a psycholinguistic perspective.
3. The author's Japanese, Level One on the Japanese-Language Proficiency Test (December, 1992), was considered sufficient for these translations.
4. It should be noted that more information might have been have been gained by looking at switching from both directions.

5. Romansh was recognized as Switzerland's fourth national (though not federally offical) language in 1938.

Ethel Ogane teaches at a private language school and a private university in the Kanto area. She has an M.A. in TESOL from Teachers College, Columbia University and is currently pursuing an Ed.D. in TESOL at Temple University Japan.

References

Appel, R., & Muysken, P. (1987). *Language contact and bilingualism*. London: Edward Arnold.

Auerbach, E.R. (1993). Reexamining English only in the ESL classroom. *TESOL Quarterly, 27,* 9-32.

Blom, J., & Gumperz, J.J. (1972). Social meaning in linguistic structure: Code-switching in Norway. In J.J. Gumperz, & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 407-434). New York: Holt, Rinehart and Winston.

Burt, S.M. (1990). External and internal conflict: Conversational code-switching and the theory of politeness. *Sociolinguistics, 19,* 21-35.

Cohen, A. D. (1987). Using verbal reports in research on language learning. In C. Faerch, & G. Kasper (Eds.), *Introspection in second language research* (pp. 82-95). Clevedon, Avon: Multingual Matters.

Faerch, C., & Kasper, G. (Eds.). (1983). *Strategies in language communication*. London: Longman.

Fotos, S.S. (1990). Japanese-English code switching in bilingual children. *JALT Journal, 12,* 75-98.

Fotos, S.S. (1994, October). Japanese-English code switching in EFL classrooms. Paper presented at the JALT International Conference, Matsuyama, Japan.

Fotos, S.S. (1995, September). Using Japanese to learn English: Codeswitching in the EFL classroom. Paper presented at the JACET (Japan Association of College English Teachers) Annual Conference, Tokyo.

Gardner-Chloros, P. (1991). *Language selection and switching in Strasbourg*. Oxford: Clarendon Press.

Gardner-Chloros, P. (1995). Codeswitching in community, regional and national repertoires: The myth of the discreteness of linguistic systems. In L. Milroy, & P. Muysken (Eds.), *One speaker, two languages: Cross disciplinary perspectives on code switching* (pp. 68-89). Cambridge: CUP.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.

Gumperz, J.J. (1982). *Discourse strategies*. Cambridge: CUP.

Kasper, G. (1994, May). A bilingual perspective on interlanguage pragmatics. Paper presented at the Tokyo-JALT Annual May Conference, Tokyo.

Kasper, G., & Blum-Kulka, S. (1993). *Interlanguage pragmatics*. New York:

OUP.

Legenhausen, L. (1991). Code-switching in learners' discourse. *IRAL, 29,* 61-73.

Lin, A.M.Y. (1988). Pedagogical and para-pedagogical levels of interaction in the classroom: A social interactional approach to the analysis of the codeswitching behaviour of a bilingual teacher in an English language lesson. *University of Hong Kong, Language Centre, Working Papers in Linguistics and Language Teaching, 11.*

Lin, A. M. Y. (1990). Teaching in two tongues: Language alternation in foreign language classrooms. *City Polytechnic of Hong Kong, Department of English, Research Report 3.*

Lipski, J. M. (1978). Code switching and the problem of bilingual competence. In M. Paradis (Ed.). *Aspects of bilingualism* (pp.250-274). Columbia, SC: Hornbeam Press.

Martin-Jones, M. (1995). Codeswitching in the classroom: Two decades of research. In L. Milroy, & P. Muysken (Eds.), *One speaker, two languages: Cross disciplinary perspectives on code switching* (pp. 90-111). Cambridge: CUP.

McClure, E. (1981). Formal and functional aspects of the codeswitched discourse of bilingual children. In R.P. Duran (Ed.), *Latino language and communicative behavior* (pp. 69-94). Norwood, NJ: Ablex.

Milroy, L., & Muysken, P. (1995). *One speaker, two languages: Cross disciplinary perspectives on codeswitching*. Cambridge: CUP.

Myers-Scotton, C. (1993a). Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society, 22,* 475-503.

Myers-Scotton, C. (1993b). *Dueling languages: Grammatical structure in codeswitching*. Oxford: Clarendon.

Myers-Scotton, C. (1993c). *Social motivations for codeswitching: Evidence from Africa*. Oxford: Clarendon.

Nishimura, M. (1989). The topic comment construction in Japanese-English code-switching. *World Englishes, 8,* 365-377.

Nishimura, M. (1995). A functional analysis of Japanese/English codeswitching. *Journal of Pragmatics, 23,* 157

Poplack, S. (1980). Sometimes I'll start a sentence in English *y termino en Español*: Toward a typology of code-switching. *Linguistics, 18,* 581-618.

Poplack, S. (1988). Contrasting patterns of codeswitching in two communities. In M. Heller (Ed.), *Codeswitching: Anthropological and sociolinguistic perspectives* (pp. 215-244). Berlin: Mouton de Gruyter.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford: Blackwell.

Scotton, C.M., & Ury, W. (1977). Bilingual strategies: The social functions of code-switching. *International Journal of the Sociology of Language, 13,* 5-20.

Tabouret-Keller, A. (1995). Conclusion: Codeswitching research as a theoretical challenge. In L. Milroy, & P. Muysken (Eds.), *One speaker, two languages: Cross disciplinary perspectives on codeswitching* (pp. 344-355).

Cambridge: CUP.

Tarone, E., Cohen, A. D., & Dumas, G. (1983). A closer look at some interlanguage terminology: A framework for communication strategies. In C. Faerch, & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 4-14). London: Longman.

(Received August 15, 1995; revised December 11, 1996)

Perspectives

Japanese and American Television Commercials: A Cultural Study with TEFL Applications

Paul Stapleton
Hokkaido University

This paper suggests that Japanese and American television commercials reflect the values of their respective cultures and are therefore pedagogically useful in the EFL classroom. A preliminary examination of 200 commercials from Japanese and American television programs indicates that there are significant differences in the types of commercials shown, with 54% of the Japanese commercials surveyed appealing to emotion, compared with only 7% of the American, and 66% of the American commercials using logic as the central theme, compared with only 10% of the Japanese. As rich sources of both cultural information and natural language, commercials provide EFL teachers with a means for promoting learner awareness of the values of the target culture through communicative language use.

本論は日本とアメリカのテレビ・コマーシャルがそれぞれの文化を反映しており、したがって英語の授業に有用であると示唆する。日本とアメリカのコマーシャル200点を調査したところ、日本では５４％が感情に訴えるものであるのに対し、アメリカでは感情に訴えるものは7%のみであり、中心的テーマとして論理を使うものが66%あった。日本では論理を使ったものは１０％のみであった。文化情報と自然な言語の豊かなリソースとしてコマーシャルは学習者にコミュニカティブな言語使用を通して目標文化の価値観に気づかせる手段となる。

Although understanding the values and beliefs of the target culture is of great assistance when learning a second/foreign language, these values and beliefs are often difficult for learners to discover. Therefore any resource which can vividly reveal cultural information should be considered valuable for teaching. This paper suggests that

BEST COPY AVAILABLE

JALT Journal, Vol. 19, No. 1, May, 1997

television commercials are one such resource and examines 200 Japanese and American television commercials to show how they can reflect the values of their respective cultures.

Culture-Based Features of Television Commercials

In the twentieth century, advertising has become a part of the cultural landscape as well as a big business. With prime time commercials in the United States costing in excess of half a million dollars for one minute (Wells, Burnett, & Mociarty, 1995, p. 379), advertisers must know the most effective means to stimulate the viewer to buy the product. This means drawing on input from such diverse fields as psychology, sociology and anthropology–essentially, the culture of the viewer (Wells et al., 1995, p. 163). Advertisers who fail to communicate effectively risk wasting millions of dollars in a matter of minutes and therefore, because such huge sums are at stake, advertisers must ensure that their product, the commercial, produces the desired result.

The Culture-Oriented Model used by advertisers in strategic planning (Wells et al., 1995, p. 752) suggests that although people have common needs, these needs are met in different ways. Thus, it is understood that a commercial which entices an American viewer to buy a product will not necessarily cause a Japanese viewer to react in the same way. An important study conducted in 1990 by Alice, a French advertising agency and IPSOS, a French research group (described in Wells et al., 1995) illustrates this point. Commercials from six European countries were presented 100 consumers from each of the same six countries. The results indicated that country-specific advertising cultures existed. Furthermore, the most competitive commercials in each market were produced by the advertisers who geared their commercial to the specific values of the local culture. Similar to rhetorical patterns (Kaplan, 1966, 1987), it is suggested that commercials reflect culture-specific values and ways of communicating.

The following section presents a classification system for commercials and an examination of television commercials in Japan and the United States. It is suggested that such analysis can reveal important differences in the way Japanese and Americans think and communicate.

The Study

Method

One hundred TV commercials from Japan and the United States respectively were video recorded during February and March, 1996. These commercials were taped from major national television networks from 18:00-23:00. The first 100 commercials recorded on each tape were used for analysis.

Regarding the classification system used to code the recorded commercials, advertising mogul David Ogilvy (1983, p. 103-110) has identified 13 commercial types commonly used in American commercials (see Table 1): humor, slice of life, testimonial, demonstration, problem-solution, talking head (the head and shoulders of a pitch-person selling the product), characters (like the underworked Maytag repairman in US commercials), rational reason why, news, emotion, celebrities, cartoons, and musical vignette. Sommers, Barnes, and Stanton (1995, p. 606) more recently used terms similar to Ogilvy's, indicating that commercials, at least in a broad sense, have not evolved beyond the types that Ogilvy outlined in the early 1980s. Wells et al. (1995, p. 493-495) has classified commercials more broadly, using terms such as storytelling or sight and sound. However, these categories contain the commercial types identified by Ogilvy, suggesting that his classification remains valid today.

Ogilvy's classification (1983) was based on American commercials, and although it was found that the 100 American commercials used in this study fit his typology, it became apparent that the nature of Japanese commercials was somewhat different. Many Japanese commercials were based on appeals to the emotions of the viewers and precise terms specifying the type of emotion appeared wanting. Thus, Ogilvy's list was modified and broken into two broad divisions to emphasize the differences between Japanese and American commercials. The first division, Logic, includes those commercials which provide rational reasons for buying a product. The second division, Emotion, includes commercials which appeal to a viewer's sentiments. As an interesting aside, Ogilvy's, firm, Ogilvy and Mather, failed in Japan both as a joint venture and on its own (Huddleston, 1990, p. 168).

Using the classification system presented below, the commercials were viewed a second time and coded into categories by the author and a second rater. This rater was an American teacher for the American commercials, and a Japanese research assistant for the Japanese commercials. Although the two raters had discussed the evaluation process, classifying commercials according to type was relatively subjective and

Table 1: Commercial Types

Logic (viewer is given a rational argument for buying the product)

1. Explanation – a voice-over describes the merits of the product
2. Slice of Life – two or more characters discuss the merits of the product
3. Testimonial – loyal users of a product testify to its virtues
4. Demonstration – a demonstration of how the product works or why it is good
5. Problem-solution – presents a problem then show how the product solves it
6. Talking head – a pitch-person extols the virtues of the product
7. Price – focuses on the value of the product
8. Comparison – compares the product to the competition

Emotion (viewer is led to associate the product with positive or high quality images)

1. Mood enhancement – creates a positive atmosphere around the product
2. Humor – attempts to amuse the viewer with images, dialogue, sounds, puns
3. Nostalgia – arouses feelings of nostalgia in tandem with the product
4. Celebrity – associates the product with a famous person
5. Musical vignette – a parade of fleeting images played to music
6. Animation – cartoons (almost always used in tandem with other methods)

(adapted from Ogilvy, 1983, pp. 103-110)

disagreement was bound to arise. For the American commercials, agreement between the two raters was 86%, with the first two dozen commercials showing the greatest disparity. Inter-rater agreement was only 75% for the Japanese commercials. This lower figure was probably due to the author's limited Japanese language ability as well as his lack of awareness of Japanese television celebrities.

An additional complication was the fact that many commercials combined elements of several types. For example, a commercial that used humor would also describe some merit of the product being advertised by means of a demonstration or explanation. Consequently, many commercials in the survey were classified as belonging to more than one type, for example, humor/demonstration, and were given two codes. For the Japanese commercials, only 19 were assigned a single code and 81 commercials were placed in two or more categories. For the American commercials, 17 were coded within a single commercial type, while the remaining 83 were assigned two or more codes. Thus, when one-way chi square tests adjusted for continuity (Hatch & Lazaraton, 1991, p. 395) were used as a test of the significance of differences in the types of Japanese and American commercials, the statistical procedure could only be used for the logic/emotion typology, where the items in each cell were independent.

Results

The results of the survey, given in Table 2, agree with previous findings regarding the culturally specific nature of television commercials ("The Enigma," 1993; Huddleston, 1990; Mosdell, 1986; Reed & Kimura, 1990). A significant difference (χ^2crit, 1 *df*, 3.84; χ^2obs. 40.06,

Table 2: Logic-Emotion breakdown
of 100 Japanese and 100 American commercials

Heading	Criteria	Japan	US
Logic	Commercials with some element of logic	42%	93%
	Commercials that were completely logic-oriented	10%	66%
Emotion	Commercials with some element of emotion	89%	34%
	Commercials that were completely emotion-oriented	54%	7%

Breakdown of types that showed cultural significance

Type	Criteria	Japan	US
Humor	Commercials with some element of humor	37%	16%
	Commercials that were completely humor-oriented	6%	1%
Celebrity	Commercials with celebrities	42%	3%
Direct Comparison	Commercials mentioning their product's competition	1%	10%
Demonstration	Commercials showing some kind of demonstration	25%	23%
Animation	Commercials with some animation	15%	6%

$p<.05$) was observed between the number of emotionally based commercials for the two cultures, with 89% of the Japanese commercials containing emotional elements, and a full 54% being completely emotional in content. Similarly, significantly more American commercials (χ^2crit, 1 *df*, 3.84; χ^2obs. 50.99, $p<.05$) were logically-oriented; 93% had some elements of logic, and 66% were entirely logical. To explain this difference, three cultural features will be discussed: group vs. individual society, intellectual roots, and the role of silence in communication.

Group versus individual society

Japan is often characterized as a group-oriented society while the United States is said to be based on the individual (Adler & Rodman, 1994, p. 51). It has been suggested that group societies are characterized by homogeneous populations with long, uninterrupted histories, this resulting in what anthropologist Edward Hall calls *high-context* communication (Hall & Hall, 1987, p. 8). Such communication tends to rely on non-verbal contextual elements to deliver a message. These include the way the words are spoken, the gestures and facial expressions that accompany the words (Tenhover, 1994, p. 86), the relationship of the speaker to the listener, and the situation in which the conversation occurs. These non-verbal features of communication are then interpreted by the listener/viewer. In fact, Hall (1983, p. 63) claims that when Japanese are confronted with the deductive methods used in a *low-context* communicative style, where the verbal element is paramount and contextual cues are of much less significance, they may feel as if someone is trying to get inside their heads to do the thinking for them. High-context communication, then, involves letting the communicators use contextual clues to process the meaning of the interaction for themselves (Adler & Rodman, 1994, p. 102).

One example of the high-context nature of Japanese communication can be seen in the thickness of *manga*, Japanese comic books. Because the context of the utterance is such an important part of Japanese communication, *manga* include a multitude of panels which set the mood of the story, taking more pages than what is common in English-language comics (McCloud, 1994, p. 80). Many of these panels are devoid of dialogue and may have no influence on the plot. Rather, they play to the reader's emotions. Comics in English tend to be much thinner because the focus is on the storyline; that is, they are goal-oriented (McCloud, 1994, p. 81).

Extending this idea into the genre of TV commercials, the significantly low number of logic-oriented commercials and high number of emotion-oriented commercials (Table 2) also suggest high-context communication. Japanese commercials tend to take an approach that avoids direct promotion of the product (Huddleston, 1990, p. 181). Instead, they use visual images and oblique references. Overt explanations that logically describe the product, if used, may be secondary. Not surprisingly, this often results in commercials that are enigmatic to foreigners ("*The Enigma*," 1993). However, such failure to comprehend can be expected by non-group members because of their lack of knowledge about Japanese culture.

Consider the following example. During the spring of 1996, a one-minute commercial for Nescafé, a brand of imported instant coffee, was seen on Japanese television. The audio portion of this commercial consisted of music interspersed with a voice-over that lasted for only 13 of the 60 seconds. The voice-over referred to the main figure, a relatively well-known Japanese artist, as a quality artist. Many of the visual sequences showed the main character walking through bamboo forests, or watching cherry blossoms fall. Only fleeting references were made to Nescafé. The viewer was meant to piece the images together to understand that Nescafé was associated with a person of high quality as well as with cultural icons such as bamboo and cherry blossoms which represent the essence of Japan. It has been suggested that Nescafé is one of the most successful foreign products promoted by Japanese advertising (Huddleston, 1990, p. 193), perhaps because of the advertisers' recognition of the Japanese preference for high-context communication.

Interestingly, three of the Japanese commercials surveyed had voice-overs which were almost entirely in English, again indicating the importance of image or mood over logical input. It has been suggested that the English language has become an important cultural symbol; it is not only a system of signs but is a sign in itself (Cheshire & Moser, 1994; Haarman, 1989), evoking the cultures where it is spoken as a native language as well as connotations associated with its use as an international language (Cheshire & Moser, 1994, p. 451). Undoubtedly, only a tiny segment of the viewership of these commercials would be able to understand the meaning of the English. The English content is not the point though; rather, it is the fact that English is being spoken. With the audio portion of these commercials incomprehensible, the average viewer is focused on the visual images. Again this is in tune with high-context communication.

Although Table 2 suggests that Japanese commercials are primarily emotion-oriented and American are logic-oriented, commercials using demonstration seem to stand out as exceptions because they are usually logic-oriented, yet both countries had similar percentages, 25% and 23%. However, demonstration commercials rely mainly on visual images and context, which is characteristic of high-context communication.

In contrast to the high context communication found in Japan, the United States is generally characterized by low-context communication (Adler & Rodman, 1994, p. 50; Hall, 1983, p. 8). Hall suggests that because Americans are members of a young country whose people have diverse backgrounds, they tend to prefer clarity and praise eloquence. With high mobility being a characteristic that dates back centuries to many Americans' European ancestry, the need to communicate with speakers of various languages has been a part of Euro-American culture and it is words alone that determine whether communication is successful (Tenhover, 1994, p. 85). Expressions like "What's the bottom line?" and "Get to the point" reveal the American desire for explicitly worded communication.

American commercials tend to reflect this low-context pattern of communication, with over 90% being logic-oriented and less than 10% being completely emotion-oriented. One American commercial viewed in this study clearly illustrates the value placed upon logical, audio input: The announcer in a car commercial could be heard gasping for air because he was talking so fast, trying to fill every second with information about the advertised car's attributes. A further example of the American desire for information over image is the "infomercial," a special type of long commercial which supplies the viewer with what appears to be a surfeit of information about a product.

Other elements suggested to characterize group culture are the expectation of conformity (Miyamoto, 1994; Mosdell, 1986, p. 19) and shared cultural knowledge and opinion. These elements are manifested in Japanese commercials through the abundant use of celebrities, who appeared in 42% compared with only 3% of the American commercials. For successful context-rich communication, it is considered important to keep the potential customer informed about the latest trends. Celebrities often act as opinion leaders who are on the cutting edge of what is in style. In this sense, celebrities are guides for the public to follow. The commercials for a number of products featuring the Los Angeles Dodgers pitcher from Japan, Hideo Nomo, in 1995 and 1996 are an example of the celebrity trend-setter genre.

On the other hand, American commercials suggest cultural values which are quite different. America describes itself as "the land of the free" and commercials such as one for Nike shoes urging viewers to "Just do it." A Ford commercial states, "We stand out." Such commercials reflect an American preference for individualism and the desire to do what one chooses without regard for what others may think. The low frequency (only 3%) of celebrities in American commercials is in keeping with Ogilvy's claim (1983, p. 109) that viewers tend to remember the celebrity but forget the product being endorsed. Furthermore, other researchers (Wells et al., 1995, p. 691) have questioned whether American viewers really believe that the celebrity uses the product.

Perhaps in an individual-oriented society people do not wish to be followers, so seeing a celebrity endorse a product is not necessarily a guarantee that the product will sell. Naturally, there are exceptions to this general rule. In recent years, market research results have shown that sport and pop icons like Michael Jordan, Michael Jackson, and Madonna have been able to transcend the celebrity taboo.

Intellectual Roots

Although both Japanese and Euro-American cultures have many and varied intellectual influences, noted scholar and former ambassador to Japan Edwin O. Reischauer (1988, p. 204) suggested that the most profound influence in Japan has been Confucianism. In contrast, American intellectual traditions have been shaped by thought from ancient Greece. Of course, such a characterization is greatly oversimplified but some values of Confucianism such as harmony and modesty are apparent in Japanese commercials, whereas manifestations of Greek philosophy in American commercials are evident in their logically-based presentation of information and their rhetorically-dominated nature.

As mentioned, almost 90% of the Japanese commercials had some element of emotion. Emotional appeals have no rational influence on the purchase of a product; instead, they attempt to touch the viewer with feelings such as humor and nostalgia which incite positive feelings towards the product. Emotionally-oriented commercials avoid boasting about a product's merits or directly comparing the product with its competition. Advertisers evidently believe that in this way competitive harmony is maintained and that the viewer is not driven to feel uncomfortable about immodest claims. In the Confucian way of thinking, claiming that one's product is better than a competitor's is not only immodest, but it is meaningless as well because it is up to the consumer to decide which product is best (Mosdell, 1986, p. 10). Interestingly, only one direct comparison commercial was found in the Japanese sample; it featured a Pepsi product opposite a Coca-Cola product. This almost complete lack of openly competitive advertising in Japan, and the frowning upon it by some authorities (Huddleston, 1990, p. 161), suggests that the Confucian ideals of harmony and modesty are exemplified in many Japanese commercials.

Although it has been noted that emotion-oriented advertising is both common and effective in the United States (Ogilvy, 1983, p. 109; Wells et al., 1995, p. 280), most of the American commercials in this study tended to be based on elements of logic and rhetoric. The need for information to support one's idea or decision is a cultural tradition derived from Greek philosophy. Plato advocated the dialectic, or debate, in which conflicting information is used in an effort to reach a conclusion. Advertisers recognize that many American viewers require such information in order to make a decision (Adler and Rodman, 1994, p. 102). This suggestion is somewhat supported by the data presented here. Ten percent of the American commercials used direct comparison, where the advertised product was presented beside its competitor and claimed to be superior, compared with only 1% of the Japanese commercials. One American commercial advertising a pain medicine actually lined up its product beside the three other market leaders, which could be identified by pill shape and logo. The pitch-person, while identifying one of the competitors, said that his product "just works better." Such immodest claims are commonly seen in American commercials. In the absence of a deep-down belief in Confucian harmony, Americans appear to respond positively to competition (Stewart and Bennett, 1991, p. 79).

Silence: In *The Silent Language* (1959), Edward Hall presented the idea that perceptions of time and space differ from culture to culture. It has been suggested (Nanda, 1994, p. 81) that Americans consider time and space as things to be filled. Time must be filled by some activity, and if space is not filled, it is thought of as empty or wasted (Nanda, 1994, p. 81). For Japanese, however, both space and time have intrinsic meaning. If there is silence during a conversation, the silence does not indicate emptiness; rather, it communicates a message.

In their use of sound, commercials can convey a sense of how a viewer understands aural messages. A common feature of Japanese commercials is their relative lack of a voice-over (Huddleston, 1990, p. 161). Instead, there is often simply music or sometimes silence. The value of silence in Japan can be related to Zen Buddhism, which can be loosely described as a way of learning which requires the student to master a skill by following formalized rituals (Hori, 1996, p. 26; Gotz, 1988). Herrigel (1953) writes that a Zen master shuns "long-winded instructions and explanations, (and the Zen teacher) contents himself with perfunctory commands and does not reckon on any questions from his pupil" (Herrigel, 1953, p. 45). Within the Zen tradition, words play only a minor role while silent communication is the essence of learning. In contrast to American culture, where stress is placed on using words to convey meaning (Stewart and Bennett, 1991, p. 157), Zen requires the student to eliminate rational thought and operate on a different plane. It should be noted, however, that although silence is an integral part of Zen, there are other explanations for the value placed on silence in Japanese culture (Matsumoto, 1988). Silence played an important role in feudal Japan, where subordinates were mostly silent in the presence of their superiors in recognition of the strict social hierarchy. Likewise, Japanese often speak of *issbin densbin* (Matsumoto, 1988, p. 48) or "heart to heart communication," where words are unnecessary because both speaker and listener are members of a tightly-knit, high-context culture and can understand non-verbal messages.

Regardless of the exact nature of its source, silence in Japan may carry deep feelings (Hall, 1983, p. 99) or may simply give listeners a chance to gather their thoughts or reflect on what has just been said. In Japanese commercials, silence is used to build a positive atmosphere around a product and Japanese viewers intuitively understand such use. In contrast, Levine and Adelman (1993, p. 71) note the American tendency to fill every space in a conversation with words. For Americans, silence is associated with awkwardness, misunderstanding, tension, shyness, disagreement (Tenhover, 1994 p. 156) or "dead air" (Hall, 1983 p. 99), all generally negative qualities. In English, silence is contrasted with eloquence, a highly desired quality since the time of the ancient Greeks. In the fourth century B.C. Isocrates described speech as follows:

> By speech we refute the wicked and praise the good. By speech we educate the ignorant and inform the wise. We regard the ability to speak properly as the best sign of intelligence, and truthful, legal and just speech is the reflection of a good and trustworthy soul . . . . If I may sum up on this subject, we shall find that nothing done with intelligence is done without speech. (Adler & Rodman, 1994, back cover)

The predominance of logic-oriented commercials with a strong rhetorical base serves as testimony to Isocrates' quote. The Japanese Nescafé commercial, whose minimal voice-over hardly mentioned the product, and the American Ford commercial, where the announcer was talking so fast his inhaling could be heard, are representative cases of the contrasting use of sound for communication in Japanese and American commercials.

Commercials as a Teaching Resource

While cultural icons such as food, sports, and even TV commercials are considered the exposed part of culture, underlying themes such as group versus individual values, intellectual roots, and silence as a form of communication are suggested to be part of culture's hidden nature (Levine and Adelman, 1993, p. xviii). This covert culture is very important for foreign language learners. However, unlike the surface or overt culture, which includes the language functions familiar to all language teachers, covert culture is more difficult to teach. It is like the proverbial iceberg, with only a small portion available for easy analysis. Teachers can explain the deeper values that govern the behavior of those who speak the target language. However, such explanations often remain inadequate because many students may be only vaguely aware of the deeper cultural values that govern their *own* behavior. Therefore, any pedagogical resource that can reveal to learners in a vivid way their own covert culture as well as the culture of the target language should be utilized.

TV commercials from the target culture serve as one such resource because of their pervasiveness, their meaning-focused use of the target language, and their rich cultural content. The data from the present study suggest that many commercials reflect important communicative values. However, most students are unlikely to be aware of the covert culture presented by commercials from their own culture, nor are they likely to be able to understand how commercials from America reflect its covert culture. Therefore, contrasting Japanese and target culture commercials as a communicative classroom activity can be illuminating to students who have only seen the surface message that a commercial displays. When introducing commercials to students at lower proficiency levels of the target language, teachers can begin by explaining some of the major cultural themes discussed above. Then they can show commercials from the learners' culture and the target culture which reflect these themes as transparently as possible. As a second step, or at higher proficiency levels, teachers can show commercials from both cultures and ask students to find deeper cultural values on their own, in an effort to develop critical thinking skills. (See the Appendix for a more detailed lesson plan.) To see familiar images in a completely new light can often leave a powerful impression on students.

Paul Stapleton teaches English at Hokkaido University.

References

Adler, R.B., & Rodman, G. (1994). *Understanding human communication*. Fort Worth, TX: Harcourt Brace.

Cheshire, J., & Moser, L-M. (1994). English as a cultural symbol: The case of advertisements in French-speaking Switzerland. *Journal of Multilingual and Multicultural Development, 15*, 451-470.

The enigma of Japanese advertising. (1993, August 14). *The Economist*, p. 61-62.

Gotz, I. L. (1988). *Zen and the art of teaching*. New York: Wilkerson Publishing Co.

Haarman, H. (1989). *Symbolic value of foreign language use*. Berlin: Moutin de Gruyter.

Hall, E.T. (1959). *The silent language*. Garden City, NY: Doubleday & Co.

Hall, E.T. (1983). *The dance of life*. New York: Anchor Books.

Hall, E.T., & Hall, M.R. (1987). *Hidden differences*. New York: Anchor Books.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.

Herrigel, E. (1953). *Zen in the art of archery*. New York: Random House.

Hori, G.V.S. (1996). Teaching and learning in the Rinzai Zen monastery. In T. Rohlen and G. LeTendre (Eds.), *Teaching and learning in Japan* (pp. 20-49). Cambridge: Cambridge University Press.

Huddleston, J. N. (1990). *Gaijin kaisha (Foreign companies)*. Tokyo: Tuttle.

Kaplan, R. (1966). Cultural thought patterns in inter-cultural education. *Language Learning, 16*, 1-20.

Kaplan, R. (1987). Cultural thought patterns revisited. In U. Conner and R. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 9-12). Reading, MA: Addison-Wesley.

Levine, D.R., & Adelman, M.B. (1993). *Beyond language*. Englewood Cliffs, NJ: Prentice-Hall Regents.

Matsumoto, M. (1988). *The unspoken way*. Tokyo: Kodansha.

McCloud, S. (1994). *Understanding comics: The invisible art*. New York: Harper Collins.

Miyamoto, M. (1994). *Straitjacket society*. Tokyo: Kodansha.

Mosdell, C. (1986). *The mirror makers*. Tokyo: Macmillan.

Nanda, S. (1994). *Cultural anthropology*. Belmont, CA: Wadsworth.

Ogilvy, D. (1983). *Ogilvy on advertising*. New York: Multimedia.

Reed, G. & Kimura, S. (1990). *CM watching*. Tokyo: Macmillan.

Reischauer, E.O. (1988). *The Japanese today*. Tokyo: Tuttle

Sommers, M., Barnes, J., & Stanton, W. (1995). *Fundamentals of marketing*. New York: McGraw-Hill Ryerson.

Stewart, E.C., & Bennett, M.J. (1991). *American cultural patterns*. Yarmouth, ME: Intercultural Press.

Tenhover, G.R. (1994). *Unlocking the Japanese business mind*. Washington, DC: Transemantics.

Wells, W., Burnett, J., & Mociarty, S. (1995). *Advertising: Principles and practice*. Englewood Cliffs, NJ: Prentice-Hall.

(Received November 3, 1996; revised December 11, 1996)

Appendix: Sample Lesson Plan for Teaching Cultural Values using TV Commercials

The following is a description of a 90-minute introductory EFL class in Japan using commercials to teach cultural values. Because many of the ideas presented are likely to be new to the students, much of the lesson is in the form of a lecture or demonstration.

Part One: As an introductory activity, the teacher asks students to write a definition of the word "culture." If this is too difficult, they can be asked to make a list of things which represent Japanese culture. After eliciting answers, the teacher draws an iceberg on the board and writes the students' answers in the appropriate sector of the iceberg, overt for the visible tip of the iceberg and covert for the submerged portion. To represent overt culture, students often refer to art, music, food, fashion, etc. The teacher should explain that covert culture, including communication style, beliefs, attitudes, values, and perceptions (Levine & Adelman, 1993, p. xviii), is the foundation upon which much of overt culture rests.

Part Two: As this is an introductory lesson, only one covert theme should be discussed. It is suggested that "group versus individual culture" is the most apparent when comparing Japanese and American commercials and, therefore, the most easily taught. Some simple examples of how group and individual cultures are manifested in daily life in Japan and the United States should be provided. For example, it could be pointed out that Japanese infants usually sleep in the same room as their parents for several years before getting their own room (Reischauer, 1988, p. 144), and so develop a stronger feeling of being a group member. American children, on the other hand, usually have their own room shortly after birth (Levine and Adelman, 1993, p. 172), a practice which instills independence from an early age.

The students then receive Worksheet 1, which presents the theme, outlines the thematic characteristics of the cultures, and lists the way in which these characteristics might be manifested in commercials. As the concepts are difficult, it is suggested that the worksheet be bilingual. The teacher reviews the contents of the worksheet with the students and discusses the concepts. The students will use this worksheet to help them understand commercials.

Part Three: The teacher shows two or three representative commercials, pointing out the cultural characteristics revealed. The students use the written information from Worksheet 2 to guide their analysis.

Part Four: The teacher shows several more commercials from each culture and the students work individually or in groups to analyze the commercials' representative characteristics, then complete Worksheet 2.

Worksheet 1

Theme	Characteristics	CM Manifestation
Group society (Japan)	high-context communication	fewer verbal explanations
	modesty	few hard sell cms
	harmony	few direct comparison cms
	village culture	no need to explain
	strict code of conformity	animation
	inspires need to escape	many celebrities
	desire to conform	(show viewers how to be part of the group)
Individual Society (the United States)	low- context communication	explanation-oriented (using words)
	focus on clarity	information oriented
	harmony is less important	boastful CMs / direct comparison CMs
	competitiveness is good	direct comparison CMs focus on price

Note: CM, from "commercial message," means "commercial" in Japanese.

Worksheet 2

Commercial	CM's Characteristics	Cultural Significance
car (US) Part 3	focus on price	competitiveness is good
	full of explanations	low-context communication
	mentions competitors	harmony is less important
coffee (Japan) Part 3	mostly silent	high-context communication
	few references to product	modesty / vagueness
	celebrity	desire to conform
Part 4:	*(Completed by students)*	*(Completed by students)*
Part 4:		
Part 4:		

Reviews

Multilingual Japan. John C. Maher and Kyoko Yashiro (Eds.). Clevedon, UK: Multilingual Matters Ltd., 1995. 164 pp.

Reviewed by
Steve McCarthy
Kagawa Junior College

This collection of papers demonstrating the actual linguistic diversity of Japan is a special double issue of the *Journal of Multilingual and Multicultural Development* (1995) in book form. It complements an earlier collection of papers describing the diversity of Japanese society (Maher & Yashiro, 1991). Now they have compiled a work more comprehensive in scope and more representative in content.

Multilingual Japan is a deceptively thin volume, dense with historical information, linguistic data, critical observations, and references for further exploration. Following the introduction come papers on the Ryukyuan, Ainu, Korean, and Chinese languages in Japan, loan words from English, returnees, and bilingualism in international families. Moreover the sole official language, Japanese, is involved with all the rest.

In the introduction, Maher and Yashiro survey the linguistic diversity of this archipelago. Yet to admit this officially would open the door to the more politically dangerous recognition of cultural diversity. They debunk the notion of "the Japanese," arguing that the inhabitants of these islands must be seen anew as just people, without the overlay of myths and stereotypes from second-hand accounts.

The authors discuss the development and standardization of the Japanese language, versus its dialects, sign language, and various minority languages. They deconstruct ideologies of monolingual-monocultural homogeneity and harmony as invented traditions. They see Nakasone's denial of the existence of minorities Japan as a symptom of the statism that, since the Meiji Era, has suppressed both minority aspirations and the sociolinguistic frameworks through which language diversity could be investigated.

The first paper on a specific language, "Ryukyuan: Past, Present and Future" by Akiko Matsumori, makes extensive references to vernacular research. "Ryukyuan" is preferred to describe the language group spoken in the formerly independent kingdom of Ryukyu, today's Okinawa Prefecture and some islands in Kagoshima Prefecture.

Matsumori details the history and geography of the Ryukyus in addition to analyzing the relationship between Ryukyuan and Japanese. Ryukyuan dialects are related to Japanese and have provided some diachronic clues in reconstructing the elusive history of Japanese. Mastumori observes that Ryukyuan is commonly called the Okinawan dialect of Japanese for reasons more political than linguistic.

In a typical case of language shift, Okinawans below the age of 40 are losing their Ryukyuan fluency and almost everyone speaks standard Japanese or rather interdialects resulting from interference during accommodation. Okinawans themselves have embraced language standardization to the detriment of local dialects, while schools have been draconian in stigmatizing non-standard Japanese usage. Ryukyuans have often been forced to change their social identity, to emphasize their common heritage with mainland Japan either in preference to American rule or for economic reasons.

"English in Japanese Society: Language within Language," by Nobuyuki Honna, does not deliver the sociolinguistic analysis promised, but does provide a valuable introduction to loan words from English. The strength of the paper is a taxonomy of seven types of borrowing patterns that involve semantic and/or structural changes. Since daily Japanese speech includes about 13 percent loan words, mostly from English, knowledge of loan words is necessary if EFL teachers in Japan are to develop strategies toward the variety of English with which the students were raised.

Honna may be listening too much to purists, though, when he writes that loan words alarm many people. An Asahi Shimbun ("*Honsha*," 1996) poll found that among the 77 percent who feel that Japanese usage is degenerating, only 6 percent cited an excess of foreign words, while 28 percent blamed youth slang. Yet Honna makes the redeeming observation that, lest people see compulsory English education as a failure, it has borne fruit by enriching the Japanese language.

"Bilingualism in International Families," by JALT Bilingualism N-SIG cofounder Masayo Yamamoto, summarizes research on bilingualism in Japan along with her own survey findings. She confines the data in this paper to families with one English and one Japanese native-speaking parent. She explains the choice of English between spouses is due to the greater English proficiency of the native Japanese speakers in most cases.

Yamamoto explains that in language use from parents to children at home, the L1 is used most often with one's children for emotional bonding or to consciously impart the language. The force of the societal language, however, is such that more Japanese is heard from the children than is spoken by the parents to them. Bi-literacy is difficult to attain, with bedtime reading by the minority language parent a necessary but not sufficient condition. Physical or linguistic conspicuousness in Japan can result in children resisting English to minimize their differences from the norm. But Yamamoto concludes, those who do become bilingual are generally admired.

Turning from the conspicuous to the partly submerged minorities, Maher and Kawanishi co-author "On Being There: Koreans in Japan." They recount the colonial history of forced labor that resulted in a million residents of Korean heritage. As Korean-medium schools were forbidden among those who stayed after the war, Korean language proficiency among the second, third, and fourth generations is rapidly declining. Today the *Soren* (North Korea-affiliated) and *Mindan* (ROK-affiliated) organizations operate school systems with a bilingual curriculum in Korean and Japanese.

As Korean schools are not accredited, 86% of Korean students attend Japanese schools in order to have any chance of entering national universities. However, many Korean-Japanese attending Japanese schools also receive materials on ethnic education which lovingly portray the culture of the homeland. One text by *Mindan* exhorts all Koreans in Japan to have ethnic consciousness, to live in dignity, and to be true internationalists *(kuk'chae'in = kokusaijin)*. Soren textbooks, on the other hand, tend to be more ideological, singing praises of the North Korean leaders. The lack of consensus among ethnic Koreans also appears in the contentious issue of maintaining Korean names or not in face of the mainstream society.

In "The Current State of the Ainu Language," Joseph DeChicchis's 164 references in several languages illustrates the extensive research on this minority and its language, both termed "Ainu." The number of officially registered Ainu is only 24,000, a result of their historical experience as a downtrodden and partially assimilated minority.

Ainu representatives' petition to the United Nations to recognize them as an indigenous minority treated unjustly, led to a bizarre government statement that the Ainu were Japanese. Similarly, Japanese scholars tend to emphasize Ainu-Japanese linguistic similarities. On the other hand, early reports creating the image of Ainu as Caucasian exaggerated their differences from Japanese. DeChicchis speculates that the Ainu language is non-Altaic but with much Altaic vocabulary, plus later borrowings from Japanese and the northern Asian languages of peoples they contacted.

John Maher then sheds light on "The *Kakyo:* Chinese in Japan." There are at least 50,000 stable residents of mostly inner-city communities using Cantonese and some Mandarin along with JSL. In the second to fourth generations there is a trend toward dominance in Japanese. Post-War *Shin-Kakyo* tend to start with Taiwanese or other dialects.

Another nearly 50,000 speakers of Chinese are neither well-established nor are they called *Kakyo*. Since 1980 there has been an influx of Mandarin-speaking students and laborers speaking various dialects, JSL, and pidgin. In addition, there are returnees from China wishing to be repatriated with their Japanese families.

Schools for *Hua-chi'iao (Kakyo)* exist in the port cities of Tokyo, Yokohama, Kobe and Nagasaki. In Yokohama with 300 pupils from kindergarten through middle school, instruction is in Chinese and Japanese by mostly Japan-born Chinese. Some Japanese children attend as well, analogous to those sent to English- or French- language international schools. At the Tokyo Chinese School, from primary school on, English further augments Mandarin Chinese and Japanese instruction. Maher thus regards the Chinese schools as a model for bilingual education in Japan.

In "Japan's Returnees," Kyoko Yashiro notes the shift from regarding returnees as problem children in need of re-acculturation to a valuable human resource in the search for internationalization. University-age returnees receive privileged quotas at many prestigious universities and enjoy an advantage in being hired by big businesses that send employees abroad. Various government and private sector organizations support them or their networking among each other. Very few returnees now have serious problems of linguistic or cultural readjustment.

However, while Japanese schools have been set up abroad to maintain L1, L2 maintenance has been neglected in Japan, particularly by public schools. Yashiro refutes each rationalization for this neglect. Her surveys of *kikokushijo* have shown that over 90% wish to maintain their L2, virtually all who have anything significant to maintain, regardless of the second language. Returnees thus warrant L2 maintenance as agents of internationalization and diversity in Japanese society.

A weakness of this collection is the lack of final editing, for which the publisher must bear some responsibility. The introduction is strident in tone and a bit disorganized in its laudable attempt to cover disparate issues not treated in the other papers. The typos may unfairly discourage readers from continuing to papers by non-native writers of English though these present relatively few obstacles to understanding.

The collection represent a reliable sociolingusitic study for scholars abroad, while language teachers in Japan can derive applications from a deeper understanding of our social context. Teachers can combat the unjust portrayal of Japanese students as a monolith and avoid blunders with submerged minorities by learning the variegated particulars beneath the ideology of sameness. For L2 pedagogy, the book suggests ways loan words could be a valuable aid to study, while for the hundreds of JALT members with international families, it provides the sketch of a road map for bilingual child-raising. The more deeply teachers are committed to Japanese society, the more useful this book will be.

References

"*Honsha teiki yoron chosa: Kokumin no ishiki wa*" ("Asahi Shimbun public opinion poll: Citizens' awareness"). (1996, January 1). *Asahi Shimbun*, p. 20.

Maher, J. & Yashiro, K. (Eds.) (1991). *Nihon no bairinguaruizumu (Japan's Bilingualism)*. Tokyo: Kenkyusha.

Oxford Advanced Learner's Dictionary, Fifth Edition, Oxford: Oxford University Press, 1995; x + 1,428 pp.

Reviewed by
Brian C. Perry
Nagoya University

The fifth edition of the *Oxford Advanced Learner's Dictionary* (OALD) is the latest work originating from A.S. Hornby's pioneering *Advanced Learners' Dictionary*, first published in Japan in 1948. Since then monolingual dictionaries, written specifically for non-native learners of English (learners' dictionaries), have become firmly established as a reference for intermediate and advanced learners. For such learners, there is general agreement that the benefits of using a dictionary in the target language more than compensate for any difficulties in understanding the definitions (Hartmann, 1992, p. 153). The fifth edition claims to have 65,000 definitions (with 2,300 being new words and meanings), 90,000 examples and 1,700 illustrated words. Included are 16 language study pages and 10 appendices. The smaller compact version, designed more for Japanese learners, differs from the standard version in that it has a soft cover and comes with headwords split into syllabic divisions.

What is different about the fifth edition is that for the first time it is based on corpora of present-day written and spoken English: the 100-million-word British National Corpus and the 40-million-word Oxford American English Corpus. The lexicographical trend towards corpus-based learners' dictionaries, which began with the *Collins* COBUILD *English Dictionary* (Sinclair, 1987), is of immense significance for two reasons. First, compilers can now scientifically analyze their data banks of contemporary language to select the most frequently used words for inclusion in their lists of, (a) headwords, and (b) defining vocabulary. Second, the example sentences–a key feature of any learners' dictionary—can now be based on those recorded as having been actually used. They can also be chosen to contain the word in question in its most frequently found collocations–and therefore be more pertinent for the user.

The layout of the OALD is clear and concise. Headwords are in large bold type with examples in italics. Idioms and phrasal verbs are also marked in bold at the end of each entry. Standard IPA phonemic transcriptions are given along with a pronunciation guide. American spellings, meanings and pronunciations are shown where required. Most learners' dictionaries emphasize their total number of entries, yet learners, unlike native speakers, tend to be more interested in the basic stock of words. Here the OALD seems to have got the balance about right. Its 65,000 entries are enough to cover the needs of almost all advanced learners whilst at the same time being reasonably compact. In comparison, the 100,000-word *Cambridge International Dictionary of English* (CIDE) (Proctor, 1995) is too heavy and bulky to be easily portable.

Hartmann (1992, p. 153) and others have summarized what the main design features a learners' dictionary should have. I will examine the OALD in the context of the three most important (adapted from Hartmann's summary), namely: 1. the definitions are geared to the more limited vocabulary of the foreign learner; 2. collocational detail is provided, usually by example sentences; and, 3. grammatical coding is detailed and specific and stylistic information is given, typically by usage labels.

Definitions: Entries in the OALD are defined using a 3,500-word core vocabulary that includes both headwords and derivatives. In contrast, a rival dictionary, the CIDE, claims a defining vocabulary of only 2,000 words. But these are only headwords, and their derivatives must be added. So, for example, one finds that *absorb* is included in both lists, but the CIDE by default also includes *absorbent, absorbency,* and *absorption*—none of which are likely to be considered common words for a learner. Furthermore, the definitions in the OALD rarely stray from the defining vocabulary so they may indeed be more restricted than other learners' dictionaries. Whether this is beneficial for the learner has been challenged by some, notably Carter (1987), who believe that limiting the vocabulary used in definitions is unnatural. Clearly there is a trade-off between accuracy and understandability. The OALD tends towards the latter.

Collocational Detail: The OALD does not, unlike the COBUILD, stick rigidly to the sentences thrown up by the corpus, but rather uses them as a basis for its examples. In this respect, the OALD comes down on the side of pedagogy in the long-running debate between those who have argued for authentic examples (Sinclair, 1987, p. xv), and those who have argued for pedagogically-based examples (Hausemann & Gorbahn, 1989, pp. 45-47). If one looks up the word *drainage* in the COBUILD we get the authentic but rather strange-sounding *line the pots with pebbles to ensure good drainage.* The corresponding example in the OALD is the short and concise *land drainage schemes.* Short phrases such as this are often found in the examples, particularly for less common words. On one hand they highlight a frequently found collocation and save space; on the other, they are too short to truly represent authentic English.

Grammatical & Stylistic Information: One of the key uses of a learners' dictionary, such as the OALD, is to help learners to encode the language for writing. It follows then that lexical information on its own would seem to be insufficient. A learner is also interested in how to put lexical units together in a way that is not only grammatically accurate, but is also stylistically appropriate. For individual entries the OALD adopts a system of coding which is clear and informative. There are 220 notes on usage under individual entries. Furthermore, a new system of simplified verb codes, (e.g. [Vnn] = verb + noun + noun) appears before the examples that illustrate them. Usage labels are precise. The entry *byte* for example, is labeled "computing" which is more detailed than "technical" or "specialized" found in the *Longman Dictionary of Contemporary English* (LDOCE) (Summers, 1995) and CIDE respectively.

It would also make sense for learners' dictionaries to include a coherent description of the grammar of English in a systematically ordered component (as called for by Lemmens & Wekker, 1991). However, the displaying of general grammar is patchy. Apart from the appendix on irregular verbs, syntactic information is contained within the misleadingly named study pages. As Walter (1996, p. 358) observes, there is little in the way of cross-referencing between them and the main body of the dictionary. If one looks up the word *idiom,* there is nothing to link it to a well-written two-page section supported by examples headed *idioms* in the study pages.

In addition, the OALD contains illustrations that are sensible and not overused, along with a number of color maps and pages presenting information on such things as the British and American Constitutions. Illustrations appear to follow the rule, "all items illustrated should be large enough to see and labeled in such a way that there is no likelihood of confusion between items" (Nesi, 1989, p. 133).

At the back there is a comprehensive set of ten appendices which take only 35 out of the 1,428 pages. They are clearly an efficient way of presenting lexical information. Take, for example, appendix 6, "Ranks in the armed forces." It would be long-winded and clumsy to fully define *captain* in the standard way. An accurate definition would have to not only state that a *captain* is a commissioned officer, but also explain its rank for each military service, and for both British and American armed forces. In tabular form one can immediately see its place in each hierarchy of ranks.

The most serious shortcoming of the OALD is the lack of information given about the frequency of usage for a particular word. Such information is especially helpful for non-native learners because it can enable them to give a word an appropriate priority for learning and production. This is particularly true for learners who are studying for exams which specify the number of English words to be mastered. It is somewhat disappointing then to find only a list of its 3,500-word defining "core vocabulary" as an appendix. The core vocabulary is of course not necessarily the same as the 3,500 most frequent words. Words such as *adjective* and *grammar* are only included for the reason that they are useful in dictionary definitions. Moreover, whether a word is in this list is not indicated in the in the main body of the dictionary. To find out, the word has to be looked up again in the list at the back. In comparison both the LDOCE and the COBUILD display the degree of frequency annotated next to the headword, with the LDOCE also showing it for both speech and writing. Looking up the word *insurance* in the OALD, one finds that it is included in the appendix of core vocabulary—the only information about frequency given. However, in the LDOCE it is marked as "S3 W2," meaning that *insurance* is ranked within the first 2-3000 words for speech but within 1-2000 in writing. The OALD is one step behind its rivals in this area.

It would be reasonable to assume that most learners require a dictionary which for them is concise, easy-to-use and above all portable. For them the fifth edition of the OALD is a first-rate learners' dictionary. Despite its shortcomings it broadly fulfills the main functions of a learners' dictionary which is to help non-native learners encode and decode language in a relevant, efficient and accurate manner. For some learners, however, the OALD will not be suitable. Those who require a weightier desktop reference would probably fare better with the CIDE. Those who attach importance to knowing the relative frequency of entries should consider the LDOCE or COBUILD. And finally, those who require a full description of the grammar of English in the same book will have to wait until a dictionary publisher has the foresight to include one.

References

Carter, R.A. (1987). *Vocabulary: Applied linguistics perspectives*. London: Allen and Unwin.

Hartmann, R.R.K. (1992). Lexicography, with particular reference to English learners' dictionaries. *Language Teaching, 25*, 151-159.

Hausemann, J.H., & Gorbahn, A. (1989). COBUILD and LDOCE II: A comparative review. *International Journal of Lexicography, 2*(1), 45-56.

Lemmens, M., & Wekker, H. (1991). On the relationship between lexis and grammar in English learners' dictionaries. *International Journal of Lexicography, 4*(1), 1-14.

Nesi, H. (1989). How many words is a picture worth? A review of illustrations in dictionaries. In M.L. Tickoo (Ed.), *Learners' dictionaries: State of the art* (Anthology Series, 23) (pp. 124-134). Singapore: SEAMEO Regional Language Centre.

Proctor, P. (Ed.). (1995). *Cambridge international dictionary of English*. Cambridge: Cambridge University Press.

Sinclair, J. (1987). *Collins cobuild English language dictionary*. London: Collins.

Summers, D. (Ed.). (1995). *Longman dictionary of contemporary English* (3rd ed.). Harlow: Longman.

Walter, C. (1996). Review of ELT dictionaries. *ELT Journal, 50*(4), 356-359.

Understanding Communication in Second Language Classrooms. Karen E. Johnson. Cambridge: Cambridge University Press, 1995. xv + 187 pp.

Reviewed by
Damian Lucantonio
The Japan Times

Understanding Communication in Second Language Classrooms is a book for teachers and teacher educators interested in the role of classroom-based research in second language education. The author, Karen E. Johnson, describes the nature of communication in second language classrooms so that more effective teaching strategies can be implemented and draws on data from actual classroom discourse to support the ideas that emerge. This allows her to provide a detailed account of the nature of communication in second language classrooms and to examine the implications of this for teaching and learning.

Her purpose is to describe how the patterns of communication are established and maintained in second language classrooms, and how these effect classroom performance, classroom learning, and language acquisition (p. 3). Johnson demonstrates that classroom communication does in fact have a distinct, if complex and variable, system. For example, the theoretical framework that underpins the book (p. 8) acknowledges the roles of both teachers and learners in the moment-to-moment activities that occur within classrooms. The meanings communicated are shaped by the actions and perceptions of teachers and students, as well as by the social, cultural, and institutional contexts within which classrooms exist. These actions and perceptions can differ widely from teacher to student, student to teacher, and student to student, which in turn can affect student participation, performance, and achievement.

The framework for understanding communication in second language classrooms used in the book is adapted from a model of communication and learning from British researcher Douglas Barnes (1976). He posits that students are active receivers of knowledge and believes that classroom learning is a negotiation between teachers' meanings and students' understandings. Furthermore, he argues that classroom communication occurs along two major dimensions. The first is the moment-to-moment actions and interactions that occur in classrooms. The second represents what teachers and students bring to the classroom in terms of knowledge and attitudes. The two dimensions interrelate to shape the classroom communication that takes place (p. 7).

The book has three main sections. Part 1, chapters 1 - 5, discusses each component of the theoretical framework and evaluates the contribution of each to classroom communication. Chapter 1 surveys the framework and illustrates the communicative demands placed on teachers and students in second language classrooms. Chapter 2 focuses on teachers' control of the patterns of classroom communication. For example, Johnson analyzes the IRE teaching sequence (the initiation act, the response act, and the evaluation act) (p. 17) which underlies the structure of many classroom lessons (Sinclair & Coulthard, 1975). The chapter examines the ways teachers use language to control the structure and content of classroom activities and the ways their frames of reference contribute to what and how they teach. Chapter 3 deals with student perceptions of the patterns of communication. It demonstrates how student perceptions of classroom activities effect both learning participation and learner performance in the classroom. Chapter 4 examines how students' culturally acquired knowledge and use of language shape their ways of talking, acting, interacting, and learning. It also explores how the language of the home and the language of the classroom can present discontinuities and problems in second language learning. Chapter 5 shows the extent to which maintaining the patterns of classroom communication can affect language learning and second language acquisition in general.

Part 2, chapters 6 - 8, uses the theoretical framework from Part 1 to examine the dynamics of communication in second language classrooms. Chapters 6 and 7 focus on teacher-student and student-student interactions and the roles these interactions play in classroom language learning and language acquisition. Chapter 8 places the framework for communication within the broader socio-cultural contexts of school-based issues and the community as a whole, and examines the effects of these factors on the second language classroom.

Part 3 (chapters 9 and 10) deals with promoting communication in second language classrooms. Chapter 9 examines how teachers can vary patterns of classroom communication to maximize learners' linguistic and interactional competencies. It also examines ways of increasing students' opportunities for language learning and language acquisition. Chapter 10, the last chapter of the book, examines how second language learners can develop their classroom communicative competence through an understanding of the nature of classroom communication.

Johnson bases her assumptions and conclusions on actual classroom discourse. Her goal is to develop more successful teaching practices by promoting more effective communication between teachers and students in second language classrooms. For example, she stresses the need for teachers to allow for greater variability in the patterns of classroom communication (p. 90). Rather than adopt a rigid and implicit approach to teaching, teachers should make the goals and structure of the class explicit to the learners and create instructional activities that allow for greater variability in both task performance and social participation. This encourages teachers to utilize a variety of methods and techniques to accommodate different learning styles and language outcomes.

The book has many strengths. First, it demonstrates that classroom communication does indeed have a system. Even allowing for spontaneity and variability, classroom communication is not a chaotic event, though second language students may perceive it as such. Second, the classroom communication system can be described using classroom data. This can lead to a better understanding of the system and to more effective teaching strategies.

Third, the book explains why learners do not always learn what teachers teach. The implications are that this usually occurs when teachers do not address the many variables involved in classroom communication that effect learning. Using classroom-based data, Johnson shows that teachers' perceptions of the "what" and the "how" to be learned in classrooms differ from those of their students. Furthermore, the perceptions differ from student to student. One role of the teachers is to be aware of these variables and to accommodate them in their teaching processes. This is the heart of the book. Suggestions for achieving it include setting up optimal conditions for learning and language acquisition (p. 87). and making explicit the socio-cultural norms underlying learning in the target language (p. 71).

Fourth, the book sheds light on why learners often learn what teachers do not teach. Johnson emphasizes that classrooms are a dynamic language learning environment. She points out that while teachers do play an important role, they are only one source of language learning. Students also learn a lot from each other. Student-student interactions have a direct impact upon classroom learning and second language acquisition (p. 128). This is an important consideration for all language teachers. She urges teachers to maximize the benefits of student-student interactions by creating opportunities for cooperative learning through group work (p. 112).

In recognizing the role that students have in classroom language learning, Johnson's research supports the findings of learner-centered approaches to teaching (Nunan 1988). Such research and teaching approaches advocate similar benefits from student-student classroom interactions for both language learning and language acquisition beyond the classroom. Researchers have found that group work fosters language acquisition by providing learners with the opportunity for productive language use and the negotiation of meaning (Nunan 1988).

Fifth, the book benefits from the many case studies from around the world. Short but specific reference is also made to Japan, for example the differences between Japanese and Anglo-American education systems and styles of learning (pp. 53-54; pp. 59-60). This directly concerns teachers in Japan and highlights the role of the socio-cultural factors that affect language learning across cultures.

A weakness of the book is its failure to recommend *how* the classroom teacher can effectively accommodate the variables that effect classroom communication. Johnson stresses the need for teachers to establish greater variability in their patterns of communication in the classroom without telling how to do this. She tells teachers what to consider in the classroom, but not how to deal with the issues that emerge. For example, Johnson provides many useful tips about the role of modeling and group work (pp. 111-114; pp. 157-158) in setting up optimal conditions for classroom learning, yet suggests few instructional activities or pedagogical techniques. Teachers might feel overwhelmed by the enormity of the task confronting them. Recognising the complexity of classroom communication is one thing, but adjusting teaching practices to deal with is another.

Another weakness of the book is its treatment of the classroom technique of modeling. Johnson stresses the importance of teachers' modeling of classroom language and classroom tasks as a way of promoting classroom communication. Yet she does not discuss the importance of modeling appropriate text-types or genres and the impact this has on effective communication (Halliday & Hasan, 1989). Research suggests the importance of providing learners with a variety of appropriate genres of language development and effective classroom communication (Martin, 1985). The role of text-type or genre language modeling is missing from the book.

Johnson provides a valuable description of classroom communication and the many patterns that effect this in second language classrooms. The book successfully combines a theoretical framework with actual classroom practice to examine a topic of interest to teacher and researchers. It raises significant issues for second language teaching, learning, and teacher training. The insights that emerge are relevant to any teacher or teacher educator interested in understanding more about the factors that influence teaching and learning in second language classrooms. It uses authentic data to encourage reflection on what actually takes place in the classrooms, rather than rely on intuition or anecodotal information. From it, teachers can draw conclusions based on sound theoretical evidence, which can help them develop more effective teaching practices. Anybody interested in clasroom-based research and its role in second language teaching will find this book of value.

References

Barnes, D. (1976). *From communication to curriculum*. Middlesex: Penguin.

Halliday, M.A.K., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.

Martin, J.R. (1985). *Factual writing: Exploring and challenging social reality*. Geelong, Victoria: Deakin University Press.

Nunan, D. (1988). *The learner-centered curriculum*. Cambridge: Cambridge University Press.

Sinclair, J., & Coulthard R.M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford: Oxford University Press.

An Introduction to Spoken Interaction. Anna-Brita Stenström. London: Longman. 1994. 238 pp.

Reviewed by
Hugh Molloy
Setagaya YMCA

This textbook/reference work (p. xiii) gives a detailed description of English conversation. Though broadly based on the 1975 model of Sinclair and Coulthard, it incorporates the work of Stenström and other writers. Stenström collected her data from the London-Lund Corpus of English conversations, a collection of about half a million words taped, transcribed, prosodically analyzed, and computerized. The English Stenström analyzes was spoken by "adult, educated British English speakers" (p. xii).

The first of the book's five chapers covers some characteristics of conversations and defines some of the common terms in conversation analysis, including *backchannels*[1] (p. 5), *pause units, tone units*, and *information units* (pp. 7-10), and the distinction between *coherence* and *cohesion* (p. 14). She introduces *adjacency pairs* (pp. 17-18) as well as the cooperative maxims of Grice (pp. 18-19), and the effect of temporal position (pp. 20-22), tonicity and pitch direction (pp. 24-25), and context (pp. 25-29) on the function of lexical items and utterances. Stenström covers much ground in

29 pages, perhaps too much for the uninitiated. Though a useful list of further readings appears at the end of each chapter, the number and frequency of specialized terms in this first chapter could be daunting to those unfamiliar with discourse analytic works. Subsequent chapters include useful exercises, but there are none here.

The second and third chapters are the heart of the book, treating interactional structure and interactional strategies. In short, Stenström explains how conversations and their parts fit together. To do this, she uses a five-level hierarchy consisting of, from largest to smallest, the *transaction*, the *exchange*, the *turn*, the *move*, and the *act* (p. 31). Other introductory texts based on or reproducing the Sinclair and Coulthard model such as Cook (1989) do not include the *turn* level. In Stenström the *act* is what the speaker wants to communicate; this is the "smallest interactive unit." A *move* is a speaker's contribution to a particular exchange. In a question-and-answer exchange a *move* would be everything comprising the answer or the question, possibly involving in an answer a preface, an answer proper, and a qualification. An *exchange* is a set of at least two moves by two different speakers. A *transaction* comprises all the *exchanges* having to do with a single topic (p. 30). Implicit in Stenström's hierarchy is a higher level, one comprising one or more *transactions,* the "conversation."

The *turn* here is "everything the current speaker says before the next speaker takes over" (p. 30). This addition does not quite fit with the levels Stenström has adopted from Sinclair and Coulthard. To be able to tell when an *act*, a *move*, an *exchange*, or a *transaction* has occurred, observers have to understand the language in which the interaction takes place. For example, if observers are unable to tell when one topic begins and the other ends, they cannot demarcate *transactions*. However, to understand when *turns* begin and end, observers do not have to understand a word of the language in which the interaction happens. They have only to be able to tell the difference between two voices.

As a convenience for tying together *transactions*, the *turn* is useful. Certainly the *turn* is convenient as a way to organize what we see in a "transcript." Whether the *turn* is a useful device for thinking about conversations as they naturally exist, as audio or audiovisual events, is another question. When we observe a conversation in time, be it on tape, on video, or live, having in hand the concepts of *transactions*, *exchanges*, *moves*, and *acts* is useful. These allow us to organize what we experience, categorize what we hear, and make some sense of the seeming mess of conversation. Furthermore, each one is a category to which we assign one or several linguistic or paralinguistic realizations in a conversation. Accumulating records of many such realizations in a given category can add to our knowledge of how language works. It does not seem as if the *turn* adds to our knowledge in this way. The concept of a *turn*, furthermore, is redundant when we observe conversations by listening or watching. As long as we can tell one voice from another, there is no need to have a special term for when one person talks. The *turn* is probably essential only in analyzing transcriptions of conversations.

Eleven pages in the second chapter are devoted to an inventory and examples of *act* types. Stenström notes three categories of *acts*: *primary acts*, which "can realize moves on their own," *secondary acts*, which "accompany and sometimes replace primary acts," and *complementary acts*, which "accompany but rarely replace primary acts" (pp. 38-39). The inventory comprises 28 primary *acts*, seven secondary *acts*, and 10 complementary *acts*.

Some of the categories seem more accommodating than others. A <statement>[2] "informs or expresses opinion" (p. 40). An <opine>, "gives one's personal opinion." An <evaluate> "judges the value of what the previous speaker said." A <react> "expresses attitude and strong feelings" (p. 39). Stenström gives one example of each act, not enough for me to be able to tell if, for example, <statement> and <opine> are to be considered different kinds of acts or if the latter is a special case of the former. Chapter 3 includes many more examples, however, and it is possible to skip back to Stenström's inventory with these in mind.

The reader should be aware that though some of Stenström's terms for *acts* are the same as those used in the Sinclair and Coulthard model, the referents for those terms are not the same in the two models. For instance, in Stenström, an <accept> "agrees to a <request>, <suggest>, etc" (p. 39), while an <accept> in the Sinclair and Coulthard model "shows [the teacher[3]] has heard correct information" (Cook, 1989, p. 47).

Stenström continues the second chapter with descriptions of *exchanges* and *transactions* and concludes with a discussion of discourse markers and interaction signals. She emphasizes that words she calls interaction signals, words such as "right," "quite," "really," and "OK" (p. 59) "can do different things in the same place" (p. 61). As examples, she offers instances of the use of "right" as *moves* such as a [respond], a [re-open], a [follow-up], as *acts*, such as a <confirm>, an <emphasizer>, and a <question> (pp. 61-62) and notes that it can furthermore function as an <appealer>, an <accept>, an <answer>, as well as a discourse marker (p. 62). This inventory of the functions is misleading. What seem to be multiple citations of one word are actually citations of several different ones. The intonation of the phonemes that make up the sounds we transcribe as "right" so affect meaning that we would be better off considering differently intoned instances as separate words. As with the case of the *move*, the work done here with "right" and other interaction signals and discourse markers would not be necessary were it not for the limitations of the recording medium. Hearing "right" intoned as a request for confirmation, as a topic closer, or as an appeal for agreement, observers would have no difficulty in immediately recognizing the function.

Much of the work done with this type of discourse analysis involves repairing the damage done by a dictionary approach wherein words meanings have been cataloged simply by orthography. From Hasan's idea of instantial lexical relations (McCarthy & Carter, 1994, pp. 203-204) and the contention that "[e]very lexical item may enter into a cohesive relation . . ." (Halliday & Hasan, 1976, p. 288), it seems clear that dictionaries often ill serve us. Dictionaries that took into account factors of intonation and formulaic language would obviate perhaps half the seeming slipperiness Stenström attacks. "Obviously, it is not possible to classify the lexical items [that function as interactional signals and discourse markers] into clearcut functional categories" (p. 67). I would disagree. It is not so obvious, unless we limit ourselves to the tools available today.

Chapter 3 is devoted to a discussion of how turn-taking works: how speakers take up, hold, and yield turns (p. 68). Especially helpful are the discussions of how *exchanges* can begin (pp. 88-109). I have found it easy to understand how conversations proceed within *exchanges*, however, how conversations move from one exchange to another has not been as obvious. Stenström points out language functions besides questions which initiate *exchanges*. Her discussion helped me recognize why my students' exchange procedures sometimes seem stilted. More importantly, it assisted me in thinking of ways to help students overcome this. To go further in understanding how exchanges proceed, however, more attention to how paralinguistic features of spoken interaction—the hems and haws—work to prepare conversant for *exchange* openings and terminations, was needed.

Chapter 4, devoted to conversation, includes a discussion of strategies for introducing and terminating topics (pp. 151-154). The focus here is akin to those of writers who take a functional-notional approach to language and language teaching such as van Ek (1975), but with greater attention paid to the strategic function of language. Stenström includes some information on how paralinguistic features, in particular pitch direction, function (see especially pp. 167-168). While the curious reader could get a more complete list of mechanisms whereby turn-taking functions from other sources such as Coulthard (1985), Stenström makes her list immediately useful by providing numerous examples.

Stenström also writes on the workings of three different categories of spoken interaction: interviews, discussions, and conversations (p. 169), examining one example of each from the London-Lund corpus. Presumably, the examples selected are typical, but as I read I found myself wondering how well other interviews, discussions, and conversations would fit her analyses.

The short, 11 page, final chapter is a discussion of discourse and grammar. Stenström analyzes several extracts from conversations in terms of her discourse units, mainly acts, tone units, and clause units, providing convincing support for a clause-, rather than sentence-, based approach to the grammar of spoken interaction.

As an introduction to one approach to the analysis of spoken language, this book serves well enough, provided the reader has outside support and keeps in mind that there are other approaches and even other manifestations of the same approach. The density of information might prove intimidating even for experienced readers, but the exercises provide a chance to assimilate the information. As a reference work for those who have some knowledge of conversational analysis, *An Introduction to Spoken Interaction* serves rather as an equivalent of a birder's field guide than as a work of ornithologic reference. This is a student's guide: The suggestions for further reading should prove useful. However, with no references in the bodies of the chapters, anyone hoping to quickly find information on particular points will be disappointed. Provided one is satisfied with the one approach used, this book is useful. It is possible, however, to question the utility of that approach for those who hope to teach how to do spoken interaction rather than to talk or think about it.

Current approaches to conversation analysis, of which Stenström's book is a good example, are a help to us as language analysts. As second language teachers, however, we need more. It is not enough for us to know what has happened in conversations. We need to know what happens in conversations as they happen. We need to know not only what conversations are like, but especially what participating in a conversation as it unfolds is like, how language is used on the fly, before the conversation is finished and ready for picking apart and analysis. How to use language is what we teach, after all. Analyzing language product does not go far enough.

Notes

1. *Italicized* terms appeared in bold in Stenström.
2. Angled brackets symbolized *acts* and square brackets symbolized *moves* in Stenström.
3. The Sinclair and Coulthard model was originally drawn up from evidence collected in a number of British primary schools (Cook, 1989, p. 46).

References

Cook, G. (1989). *Discourse*. Oxford: Oxford University Press.

Coulthard, M. (1985). *An introduction to discourse analysis* (2nd ed.). London: Longman.

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

McCarthy, M., & Carter, R. (1994). *Language as discourse: Perspectives for language teaching*. London: Longman.

Sinclair, J., & Coulthard, R.M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford: Oxford University Press.

van Ek, J. A. (1975). The threshold level. In C. J. Brumfit & K. Johnson (Eds.), (1979). *The communicative approach to language teaching* (pp. 103-116). Oxford: Oxford University Press.

A Communicative Grammar of English, 2nd Edition. Geoffrey Leech & Jan Svartvik. Harlow, Essex: Longman Group Limited, 1994. 423 pp.

Reviewed by
Kevin Varden
Meiji Gakuin University

One of the most difficult things for an English learner to master is understanding the use of intonation to convey meaning. Grammar references are usually of little help, since they restrict themselves to the *syntax* of English: word order, collocations, verb agreement, among others. Enter the most useful aspect of *A Communicative Grammar of English*. This is not to categorize the work as only a reference of intonation—it by no means ignores traditional grammatical information, it simply adds another welcome dimension. (For examples of a third grammar of paralinguistic gestures, see Clark, Moran, & Burrows, 1981, pp. 253-270).

The authors have contributed to two other recent grammars, *A Student's Grammar of the English Language* (Greenbaum & Quirk, 1990) and *A Comprehensive Grammar of the English Language* (Quirk, Greenbaum, Leech & Svartvik, 1985). However, the organization and communicative focus of this grammar make it more than simply a condensed version.

Changes in the Second Edition

Users familiar with the first edition's format will have no difficulty using this edition. The separate entry format for each topic remains, although the total number of entries has been reduced from 886 to 747 and the number of pages has grown from 324 to 423.

Reorganization is evident in the Table of Contents. Instead of listing Part One only as "Varieties of English" the authors retitle it more descriptively as "A Guide to the Use of this Book" with individual listings of each entry. A shortened "Note on Phonetic Symbols" is entered here. The authors include and expand entries previously in Part Two of the first edition, including "Intonation" (a general discussion of its uses). Part Three, "Grammar in Use," has become Part Two, and Part Four, "Grammatical Compendium," has become Part Three with the more student-friendly title "A-Z in English Grammar." Of note is the deletion of the decidedly student-unfriendly use of abbreviations in the discussions of verb patterns (e.g., T1 for a transitive verb with a single object), along with the somewhat cryptic table explaining them.

Another change is the introduction of corpora data, following the trend of newer grammars such as Collins COBUILD (Sinclair, 1990). This provides the user with uncontrived examples of English usage, a welcome edition. As a result, where appropriate, examples contain the errors, hesitations, and false starts that can not be avoided in normal speech. Separate entries discussing these have been added to assist the user. However, only a portion of the examples are new; many examples remain unchanged from the first edition.

Like the first, the second edition contains cross-references to a larger grammar, in this edition to *A Comprehensive Grammar of the English Language* (Quirk, Greenbaum, Leech & Svartvik, 1985). In keeping with the times, the authors have removed the generic use of the male pronoun in examples and included a separate discussion on non-sexist usage. All in all, the grammar has been given an attractive facelift. Whether or not this justifies spending the money to replace a first edition is questionable since much of the information it contains has not changed.

Contents

The main parts of the grammar, "A Guide to the Use of this Book," "Grammar in Use," and "A-Z in English Grammar," are preceded by a "List of Symbols" used throughout and followed by the index. There are no references, recommended further readings, or corpus acknowledgments, a minor failing that could have been easily remedied.

The symbols and abbreviations used are for the most part standard; neither instructors nor students should experience any difficulty using them. The use of angled brackets to indicate the variety of English (<written>, <spoken>, <formal>, <informal>, <BrE> for British English, <AmE> for American English, etc.) may be unfamiliar, as may the use of single and double vertical bars (| and ||) to delineate tone unit boundaries and separate dialectal differences respectively. When making points about intonation, raised and lowered text help clarify the intonational contours. In addition, the cuisenaire-rod style symbols used in the discussion of tense and aspect are much more helpful than simply describing them in words.

The work contains excellent discussions on many of the techniques speakers use when communicating orally. This includes the use of many filler and repair words, a refreshing inclusion—their use may not be grammatical, but they cannot be avoided in real-world oral communication. The authors included written communication in the definition of the Communicative Method, and so the text contains many references to literary communicative techniques. They discuss literary style in a separate entry as well as in notes throughout. There are also entries on similarities and differences between the two modes of communication.

The text differentiates between British and American English via the abbreviations <BrE> and <AmE>. Other dialects of English are not represented, since ". . . the varieties of English used in the United States and in Britain are the most important in terms of population and use throughout the world . . ." (p. 28). Whether this position is defensible or not, and whether or not other dialects contain noteworthy grammatical differences, as opposed to differences in lexicon or pronunciation, I do not have the experience to say.

Part 1: "A Guide to the Use of this Book," contains information necessary to use the grammar as well as the background information needed to understand the entries. This includes discussions of the inherent differences between oral and written modes of communication; interactive, non-interactive, and cooperative uses; characteristics of spoken English; intonation; the use of phonetic symbols; geographical differences, and the uses of formality, politeness, and tactfulness. All in all, quite a useful summary of the use of the English language in itself.

Linguistically, the discussion on intonation is simplistic in the use and meanings of the three tonal patterns it distinguishes, but it is a thorough introduction nonetheless. Examples are typically marked for nuclear stress, tone, and intonational group; the use of raised and lowered text mentioned earlier aids understanding. A great deal of information is presented in a manner accessible to a Japanese college student.

Part 2 contains the main grammar section. There is a wealth of information contained in this section, called "the central part and the largest part of the grammar" by the authors (p. 6). It contains four sections in itself, and can be visualized as four concentric circles, each adding a layer of complexity to the communicative act.

Section A: "Concepts," represented by the innermost circle, contains discussions of concepts associated with our experience in the world. These include notions of nominal reference, abstractness, quantity, definiteness, relations between ideas expressed by nouns, restrictiveness, tense and aspect, duration, frequency, spatial relations, causality, condition, degree, role, comparison, inclusiveness, and topic marking. In particular, the diagrams that accompany discussions of nominal reference, quantity, degree, and spatial relations are small but clear, and quite useful.

Section B: "Information, Reality, and Belief," contains discussion of the giving and receiving of information within the context of the logic that we continually process information with. Entries here include statements, questions and responses; omission of information; reported statements and questions; denial and affirmation; agreement and disagreement; fact, hypothesis, and neutrality; degrees of likelihood; and attitudes of truth. The heavy use of stress marking in many examples accurately reflects the use of intonation in providing non-lexical information.

Section C: "Mood, Emotion, and Attitude," discusses the pragmatic use of the language in achieving a given task. This includes emotive emphasis in speech, describing emotions, volition (conscious choice), permission and obligation, influencing people, friendly communications, and vocatives (forms of address). Again, since intonation is crucial in the display and interpretation of moods, emotions and attitudes, the authors include stress-marked examples.

Finally, Section D: "Meanings in Connected Discourse, " represented by the outer-most circle, discusses the context sensitivity of sentences to surrounding material—the place of an utterance in the communication as a whole. Topics are linking signals, linking clauses and sentences, "general purpose" links, cross-reference and omission, presenting and focusing information, and order and emphasis. The entries under "cross-reference and omission" should be particularly helpful to Japanese students, since English "reference at a distance" is often problematic. An entry on reference to understood knowledge, another major source of non-understanding is missing and would improve the grammar. (Try asking your students to explain the slogan "Budweiser: It's it, and that's that.")

Part 3: "A-Z in English Grammar," resembles a more traditional grammar reference, with the addition of intonational marking on various examples. It provides the user with an alphabetically arranged discussion of many grammatical topics. There are, among many other things, an excellent discussion on the three groups of determiners (grouped by position in a noun phrase), an extensive listing of irregular verbs both alphabetically and by group, and a thorough listing of nationality words. This section also contains a large number of cross-references, to both entries within this text and in the larger *A Comprehensive Grammar of the English Language* (Quirk et al., 1985).

Several typefaces are utilized in the index: italics for individual words, all capitals for grammatical terms, plain text for functions and meanings, and angled brackets to indicate entries on variety. Thoughtfulness in preparation is evident; for example, although the authors use the term *concordance* in the text instead of *agreement*, there are entries for both in the index.

The book is very free from errors. In consulting it extensively, I only came across one cross-referencing error—the discussion of willingness (entry 320, p. 161) refers the reader back to entry 129 on uses of *will*; this entry is actually entry 140. I disagree with the continued usage of the traditional label *verb phrase* (p. 396) to refer to combinations of verbs (i.e., main verbs and their associated modals and auxiliaries) since the term is now ambiguous between this usage and the broader definition of the whole of the predicate used in generative grammar.

These minor caveats aside, this is a reference book that I both use and refer to my students. It provides much of the information students need to continue developing their communicative ability. Instructors will find it an excellent resource as well. In particular, the inclusion of intonation is a great improvement over standard grammars. I recommend this work to Japanese students over either of the two larger works the authors have contributed to; while the grammatical information contained in them is more detailed, the communicative focus of this work makes it more applicable to the needs of the average student. However, whether one would want to spend the money to replace a first edition of this work would probably depend on how worn out the covers are.

References

Clark, R.C., Moran, P.R., & Burrows, A.A. (1981). *The ESL miscellany*. Holt-Saunders International Edition. Tokyo: Holt-Saunders Japan Ltd.

Greenbaum, S., & Quirk, R. (1990). *A student's grammar of the English language*. Harlow, Essex: Longman.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English*. Harlow, Essex: Longman Group Limited.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Harlow, Essex: Longman.

Sinclair, J. (1990). *Collins cobuild English grammar*. London: Harper Collins Publishers.

Functional English Grammar: An Introduction for Second Language Teachers. Graham Lock. Cambridge: Cambridge University Press, 1966. 296 pp.

Reviewed by
Randall Davis
Nagoya City University

As part of the Cambridge Language Education Series on teaching, *Functional Grammar* attempts "to provide teachers with ways of thinking about English grammar" and to aid us in "understanding the kinds of problems learner may have" associated with it (p. xi). However, while Lock's aim is admirable, he fails to provide a text that is within the linguistic and practical reach of layman readers, presumably ESL teachers, which is something the title and preface clearly suggest.

The back cover entices the reader saying that the book has been written to cover "areas of difficulty for second language learners" by providing examples, tasks, and teaching discussion questions to explore these particular problems. Chapter 1, "Some Basic Concepts" defines functional grammar, outlines how grammar is used to express meaning, and defines a number of basic "basic" terms. The next eleven chapters attempt to examine grammar and how it is used, including chapters such as "Doing and Happening I: Transitivity of Action Processes" (Chapter 4), "Doing and Happening II: Ergativity, Phrasal Verbs, and Phrase" (Chapter 5), and "Representing Time: Tense and Temporal Adjuncts" (Chapter 8). With such complex topics, the book is in fact better as a classroom text on grammatical structure for linguistics than for self-study for classroom teachers looking for an introductory guide to English grammar.

To be fair, the latter half of the book is a little easier to read and more interesting with its focus on different kinds of speech acts, expressions of attitude, opinions and judgments, and textual meaning, all as part of functional grammar. Even then, it is heavy reading, which tends to alienate the "introductory" reader and does little to dispel the criticisms against grammar teaching which Lock discusses in Chapter 13.

Ironically, the most informative and thought-provoking part is this chapter "Issues in the Learning and Teaching of Grammar." In his overview of the history of grammar teaching, Lock notes that although a structural grammar methodology "has tended to be associated with a more or less discredited approach to second language teaching," there is a resurgence of interest in the area. He goes on to point out several methodological options for teaching language, emphasizing that the teaching of grammar needs to be interwoven into the instruction of the other language skills through "meaningful and motivating activities" (p. 277) that provide opportunities for students to solve problems by creating generalizations about how grammar works. One idea that Lock refers to is an information transfer technique where students transfer spoken or written input to charts or graphs (e.g., for a presentation) and thus expose themselves to certain grammatical features.

Unfortunately, such moments are few and far between. *Functional English Grammar* turns out to be a reference tool appropriate for the researcher or academic linguistic, not the introductory guide to grammar teaching teachers want to help them class. The promises contained in the subtitle and on the covers, that readers need "no previous study of English grammar or linguistics" are unmet as readers wade through chapter after chapter trying to decipher the complex explanations of "functional" grammar. To end on a bright note: readers looking for a truly "functional" book on grammar can find several listed in the references that do fit the bill.

Spoken English on Computer: Transcription, Mark-up and Application. Edited by Geoffrey Leech, Greg Myers, and Jenny Thomas. Essex, England: Longman Group Limited, 1995. 260 pages.

Reviewed by
Charles Adamson
Miyagi University

Corpora, large bodies of language data (often annotated) stored in computer readable form, are becoming increasingly important. Linguists are using them to investigate language as it actually is used, rather than as they think it might be. At conferences and workshops, there are frequent references to corpora. Recently textbook writers and classroom teachers have begun deciding what should be taught on the basis of data from corpora. It appears that this trend will continue and a knowledge of corpora and their uses will be a necessary part of a TEFL/TESL teacher's education and hence a matter of interest for JALT members.

Spoken English on Computer (SEOC) is a good introduction to the field and a must read for anyone working with any of the established corpora of spoken language or considering the development of a new one. The book began as a collection of papers that were presented at a *Workshop on Computerized Spoken Discourse* held in England in September, 1993. The editors, however, did not feel that the papers covered the field sufficiently so they elicited papers from other workers in the field. SEOC is the result.

Each of the book's three sections has an introduction by the editors: Part A, with seven articles addressing fundamental issues; Part B, with seven articles discussing the application of corpora of spoken language; and Part C, with six articles describing corpora which exist or are being constructed. While the subject matter is quite technical, the authors make the content accessible to non-specialists.

The first four chapters of Part A analyze the theory and principles of transcription. The most problematic area is what data should be included in the corpus and what form that data should take. This applies especially to para-linguistic information. The next two chapters deal with the recommendations of the Text Encoding Initiative (TEI) and their implications. The TEI is the result of joint development by a number of groups and outlines a markup method based on Standard Generalized Mark-up Language (SGML). While this treatment is insufficient to allow the reader to actually mark up texts, it is an excellent introduction and would be of great value to any teacher or researcher using a corpus that uses TEI. In the final chapter of the section, John Sinclair discusses economical ways to construct corpora of spoken language, the need for better definitions of "spoken material" and the need to make corpora more user-friendly.

Part B balances the theory from Part A with practical examples of the use of spoken language corpora including the study of language impairment, code-switching, and the structure of conversation. There are also discussions of the tagging of time, intonation, and word class. Another chapter describes a research project in which conversations were recorded while native speakers were doing information gap activities on maps. The editors point out that these chapters when taken together raise a number of practical problems: the quality of the original recordings; confidentiality for the sources; the importance of information that is lost through transcription, coding, and mark-up; and the hardware that is required.

Part C contains descriptions of six corpora of spoken language: (1) The London-Lund Corpus; (2) The COBUILD spoken corpus; (3) The Machine Readable Spoken English Corpus (MARSEC); (4)The International Corpus of English; (5) The BNC spoken corpus; and (6) The Bergen Corpus of London Teenager Language (COLT).

More than 500 references have been gathered together in a Bibliographical References section at the end of the book along with an Author Index and a Subject Index. The latter, however, is not particularly useful since it has only about 120 entries.

In the General Introduction the editors indicated that various strands of research are currently coming together in the field of spoken language corpora. This book, which provides chapters discussing some of the important aspects of many of these strands, is an excellent overview or introduction to the field for language teachers and researchers in Japan.

Books to Review

Reviews are being sought of the following books. Contact the Reviews Editor (see Guidelines) for further information.

Baynham, M. (1995). *Literary practices: Investigating literacy in social contexts.*

Brown, J.D. (1996). *Testing in language programs.*

Durand, J., & Katamba, F. (1995). *Frontier of phonology: Atoms, structures, derivations.*

Freeman, D., & Richards, J.C. (1996). *Teacher learning in language teaching.*

Grabe, W., & Kaplan, R.B. (1996). *Theory and practice of writing.*

Gutknecht, C., & Röle, L.J. (1996). *Translating by factors.*

Hasan, R., & Williams, G. (Eds.). (1996). *Literacy in society.*

Hatch, E., & Brown, C. (1995). *Vocabulary, semantics, and language education.*

Jacobs, R.A. (1995). *English syntax: A grammar for English language professionals.*

Kunnan, A.J. (1995). *Test taker characteristics and test performance: A structural modeling approach.*

McKay, S.L., & Hornberger, N.H. (eds.) (1996). *Sociolinguistics and language teaching.*

Nunan, D., & Lamb, C. (1996). *The self-directed teacher: Managing the learning process.*

Pennington, M.C. (1996). *Phonology in English language teaching: An international approach.*

Quirk, R. (1995). *Grammatical and lexical variance in English.*

Ridout, R.M. (ed.) (1996). *The Newbury Hourse dictionary of American English: An essential reference for learners of American English and culture.*

Sarangi, S., & Slembrouck, S. (1996). *Language, bureaucracy and social control.*

Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses.*

Spada, N., & Frölich, M. (1995). *COLT (Communicative orientation of language teaching) observation scheme: Coding, conventions and applications.*

Street, B.V. (1995). *Social literacies: Critical approaches to literacy in development, ethnography and education.*

Swan, M. (1995). *Practical English usage (2nd ed.).*

Thomas, J., & Short, T. (Eds.). (1996). Using *corpora for language research.*

van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity.*

Wodak, R. (1996). *Disorders of discourse.*

Information for Contributors

Contributions must conform to the *JALT Journal* Editorial Policy and Guidelines.

Editorial Policy

JALT Journal, the research journal of the Japan Association for Language Teaching (*Zenkoku Gogaku Kyoiku Gakkai*), welcomes practical and theoretical articles concerned with foreign language teaching and learning in Japanese, Asian, and international contexts. It welcomes well written contributions which provide readers with a link between theoretical and practical issues, especially those addressing current concerns in pedagogy, methods, and applied linguistics. Areas of specific interest are:

1. curriculum and teaching methods
2. classroom centered research
3. cross-cultural studies
4. teacher training
5. language learning and acquisition
6. overviews of research and practice in related fields

The editors encourage submission of full-length articles, short research reports, book and media reviews, essays on trends in language education, reports of pedagogical techniques which are thoroughly framed in theory and supported by descriptive or empirical data, and comments on previously published *JALT Journal* articles. Occasionally, *JALT Journal* will issue a Call for Papers for theme-based issues. Articles should be written with a general audience of language educators in mind, with statistical techniques and unfamiliar terms clearly explained or defined.

Guidelines

Style

JALT Journal uses the *Publication Manual of the American Psychological Association* (available from the Order Department, A.P.A., 1200 17th St. N.W., Washington DC). Consult recent copies of *JALT Journal* or *TESOL Quarterly* for examples of documentation and references. Give the page numbers of cited works in both the text and references.

Format

No more than 20 pages (approximately 6000 words) for full-length articles, including reference list, typed and double-spaced on A4 or 8.5"x11" paper. *Research Forum* submissions should be no more than 10 pages. *Perspectives* submissions should be no more than 15 pages. Authors must supply camera-ready diagrams or figures (if any) before final publication. The author's name and references that identify the author should appear only on the cover sheet.

Materials to be submitted

Three (3) copies of the manuscript
Cover sheet with title, running head title (one to four words), and author name(s)
Abstract (no more than 150 words)
Japanese translation of title and abstract if possible (less than 400-*ji*)
Biographical sketch(es) (no more than 50 words)
Authors of accepted manuscripts must supply a disk copy (RTF or ASCII).

Evaluation Procedures

All manuscripts are first reviewed by the editorial staff to insure they comply with *JALT Journal* guidelines. Those considered for publication are subject to blind review by at least two readers, with special attention given to the *JALT Journal* Editorial Policy, the significance and originality of the submission, and the use of appropriate research design and methodology. Evaluation is usually completed within three months.

Restrictions

Papers submitted to *JALT Journal* must not have been previously published, nor should they be under consideration for publication elsewhere. *JALT Journal* has First World Publication Rights, as defined by International Copyright Conventions, for all manuscripts published. We regret that manuscripts or computer disks cannot be returned.

Address for Manuscripts and Inquiries

Tamara Swenson, *JALT Journal* Editor
Osaka Jogakuin Junior College
2-26-54 Tamatsukuri, Chuo-ku, Osaka 540 Japan

Submissions to *Perspectives*

The editors invite reports of pedagogical techniques which are thoroughly framed in theory and supported by descriptive or empirical data to the *Perspectives* section. Send submissions to:

Sandra Fotos, *JALT Journal* Associate Editor
Senshu University, 2-1-1 Higashi Mita
Tama-ku, Kawasaki, Kanagawa 214 Japan

Reviews and Address for Reviews

The editors invite reviews of books, tests, teaching systems, and other substantial publications in the field of language education. A list of publications which have been sent to JALT for review is published monthly in *The Language Teacher*. Please send requests for books, materials, reviews guidelines, and queries to submit reviews to:

Thomas Hardy, *JALT Journal* Reviews Editor
Tamagawa University, 6-1-1 Tamagawa Gakuen
Machida-shi, Tokyo 194 Japan

Japanese-Language Manuscripts

The *JALT Journal* welcomes contributions on teaching Japanese as a second language (JSL). Submissions must conform to the above Guidelines. Authors are requested to proved a detailed abstract in English, 500-750 words in length. Refer to the Japanese-language Guidelines for details. Send all Japanese-language manuscripts to:

Naoko Aoki, *JALT Journal* Japanese-language Editor
Shizuoka University, Department of Education
836 Oya, Shizuoka 422 Japan

Address for Inquiries about Subscriptions or Advertising

JALT Central Office
Urban Edge Bldg. 5F
1-37-9 Taito, Taito-ku, Tokyo 110 Japan
Tel: 03-3837-1630; Fax: 03-3837-1631
(From outside Japan: Tel: 81-3-3837-1630; Fax: 81-3-3837-1631)

JALT Journal 日本語論文投稿要領

JALT Journal は、日本語教育に関する日本語論文の投稿を受け付けます。第二言語教育／学習研究の文脈を踏まえ、しっかりした研究計画に基づいている、実践への応用の可能性を示した理論的・実証的研究や理論的裏づけを持った実践報告などで、教育、教授法、応用言語学などの今日的問題を扱ったものを歓迎します。

文体は一般的な学術論文のスタイルを使ってください。章だてのしかたや参考文献のデータの書き方などは、Publication Manual of the American Psychological Association (4th. ed.) の定める方式にできるだけ近い形にしてください。ご不明の場合は、JALT Journal の英語論文を参考にするか、日本語編集者までお問い合わせください。また、JALT Journal の読者は現場の教師が主ですから、特殊な専門用語や統計的手法は、わかりやすいように定義したり説明したりしてください。

原稿の長さは、参考文献リストも含め400字づめ原稿用紙30枚以内です。A4の用紙に横書きで、1行40字、1ページ30行で印刷してください。手書きの原稿は受け付けられません。図表を入れる場合は、JALT Journal のページのサイズに合わせて作成してください(縮小コピー可)。執筆者の名前や所属機関名は、原稿には書かないでください。ご提出いただくものは、以下の通りです。

- 原稿3部
- 執筆者の名前と所属機関名を書いた表紙
- 400字以内の和文要旨
- 英文のタイトルと、500～750語の英文要旨
  英文要旨が用意できない場合は、400字づめ原稿用紙3枚程度の詳細な和文要旨を
  ご提出ください。
- 100字以内の執筆者略歴
- 可能ならば、Macintosh のテキスト形式で保存したファイル

査読の方法は、編集委員会で投稿要領にあっているかどうかを確認したあと、少なくとも二人の査読者が査読を行います。査読者には執筆者の名前は知らされません。査読の過程では特に、原稿がＪＡＬＴ Journal の目的にかなっているか、言語教育にとって意味があるか、独創性はあるか、研究計画と方法論が適切かが評価されます。査読は通常二か月以内に終了します。

JALT Journal に投稿する原稿は、すでに出版されているものや他の学術雑誌に投稿中のものは避けてください。JALT Journal は、そこに掲載されるすべての論文に関して国際著作権協定による世界初版権をもちます。

お送りいただいた原稿は返却できません。

投稿原稿の送り先とお問い合わせは以下にお願いします。

〒422　静岡市大谷836　静岡大学教育学部
JALT Journal 日本語編集者　青木直子

JALT Journal 第19巻　第1号

1997年5月　6日　印刷

1997年5月15日　発行

編集人　タマラ・スウェンソン
発行人　ジーン・ヴァン・トロイヤー
発行所　全国語学教育学会事務局
〒110東京都台東区台東１－３７－９アーバンエジビル５Ｆ
TEL (03)3837-1630; FAX (03)3837-1631
印刷所　コーシンシャ株式会社
〒530 大阪市北区天満2-13-3 TEL (06)351-8795

The 23rd Annual
International Conference
on Language Teaching /
Learning & Educational
Materials Exposition

JALT 97

Oct. 9 - 12

Act City
Hamamatsu, Japan

Theme:
Trends & Transitions

FOR MORE INFO CONTACT:
JALT Central Office
Urban Edge Bldg. 5th Fl
1-37-9 Taito, Taito-ku, Tokyo 110
Tel: 03-3837-1630 FAX: 03-3837-1631

BEST COPY AVAILABLE

Visit the JALT97 WWW Homepage at:
http://www.miyazaki-mic.ac.jp/JALT/JALT97.html

Leaders in Linguistics

Analysing Sentences: An Introduction to English Syntax, SECOND EDITION
Noel Burton-Roberts
Ppr 0 582 24876 0 328pp Due mid '97
LEARNING ABOUT LANGUAGE SERIES

The Construction of Professional Discourse
Edited by Britt-Louise Gunnarsson, Per Linell & Bengt Nordber
Ppr 0 582 25941 X Csd 0 582 25942 8 344pp '97
LANGUAGE IN SOCIAL LIFE SERIES

Adverbs and Modality in English
Leo Hoye
Ppr 0 582 21535 8 336pp '97
ENGLISH LANGUAGE SERIES

Creating Texts: An Introduction to the Study of Composition
Walter Nash and David Stacey
Ppr 0 582 24486 2 256pp '97
ENGLISH LANGUAGE SERIES

Language, Literature and the Learner: Creative Classroom Practice
Edited by Ronald Carter & John McRae
Ppr 0 582 29323 5 256pp '96
APPLIED LINGUISTICS AND LANGUAGE STUDY SERIES

Language and Development: Teachers in a Changing World
Edited by Brian Kenny & William Savage
Ppr 0 582 25866 9 384pp '96
APPLIED LINGUISTICS AND LANGUAGE STUDY SERIES

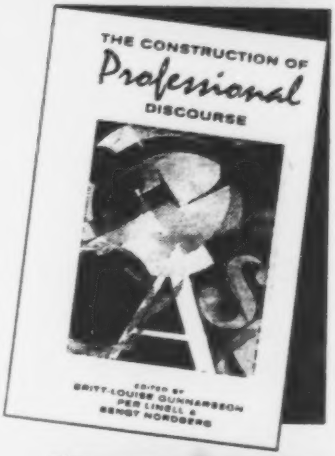
Autonomy and Independence in Language Learning
Edited by Phillip Benson & Peter Voller
Ppr 0 582 28992 0 288pp '96
APPLIED LINGUISTICS AND LANGUAGE STUDY SERIES

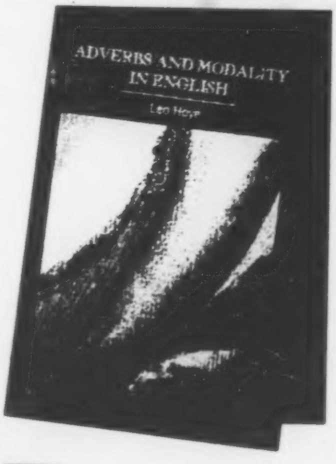
Language, Bureaucracy and Social Control
Srikant Sarangi & Stefaan Slembrouck
Ppr 0 582 08622 1 Csd 0 582 08623 X 256pp '96
REAL LANGUAGE SERIES

For Further information on these and other Longman titles, contact Mieko Sumi

Longman Japan, Gyokuroen Building
1-13-19 Sekiguchi, Bunkyo-ku, Tokyo 112
Tel: 03-3266-0404; Fax: 03-3266-0326

ANALYSING SENTENCES





LONGMAN

第二十一巻 第七号 平成九年五月十五日発行 〒11 東京都台東区台東一—三七—九 アーバンエッジビル五階 デビット・マクマレー
昭和五十四年四月二十一日 第三種郵便物認可
THE LANGUAGE TEACHER 特別増刊号

BEST COPY AVAILABLE

SSN 0287-2420 ¥950

全国語学教育学会

JALT JOURNAL

November 1997
Volume 19 • No. 2
Japan Association for Language Teaching

BEST COPY AVAILABLE

Inside this issue:

Questionnaire validation
Reading-writing • Music
Reflective journals
Course design • Empathy
Student responses • Haiku

Nellie's
DISCOUNT BOOKS

The ELT Specialists

English Language Teaching

At Nellie's we deliver the goods! Whether you need one text or a thousand. you won't find anyone who'll get them to you faster or cheaper than Nellie's. Just phone or fax us for your free catalogue, and our NEW toys and video catalogue.

LTP "A Reputation For Innovation"

As part of our constant efforts to improve Nellie's services even more we have become the sole distributor in Japan for LTP, an innovative independent publisher specialising in business and conversation texts influenced by "The Lexical Approach". For more information see Nellie's 1998 catalogue, or contact us for an LTP catalogue

For free catalogue and information call
0120-071329 (Free Dial)
Fax: 03-3676-1815 e-mail nellies@jms.jeton.or.jp

BEST COPY AVAILABLE

Contents

November 1997
Volume 19 • No. 2

全国語学教育学会
JALT
JOURNAL

173 In this issue
175 From the Editors

Articles

177 Validating a Questionnaire on Confidence in Speaking English as a Foreign Language
Dale T. Griffee
198 On Reading-Writing Relationships in First and Foreign Languages
Ahmad Abu-Akel
217 Teaching with Music: A Comparison of Conventional Listening Exercises with Pop Song Gap-fill Exercises
Kim Kanel
235 Learning Pronunciation and Intonation of Japanese through Drama by Beginning Language Students: A Case for Reflective Journals
Harumi Moore
260 オーストラリアにおける観光業用の日本語コースのデザイン
(Course Design and Delivery of Hospitality Japanese in Australia)
Fusako Osho, Hiromi Masumi-So, & Chihiro Kinoshita Thompson

Perspectives

271 Empathy and Teacher Development
John B. Kemp
279 Answer, Please Answer! A Perspective on Japanese University Students' Silent Response to Questions
Timothy J. Korst
292 A Poem in the Process: Haiku as an Alternative to Brainstorming
John Esposito

Reviews

309 *Teachers as Course Developers* (Kathleen Graves, Ed.) – Reviewed by Greta J. Gorsuch
312 *Verbal Hygiene* (Deborah Cameron) – Reviewed by Virginia LoCastro
318 *Understanding Metaphor in Literature: An Empirical Approach* (Gerard Steen) – Reviewed by Valerie Fox
324 *The Functional Analysis of English: A Hallidayan Approach* (Thomas Bloor and Meriel Bloor) – Reviewed by Wendy L. Bowcher
327 *Teaching Business English* (Mark Ellis and Christine Johnson) – Reviewed by Steffen Eckart

329 *Literacy and Script Reform in Occupation Japan: Reading Between the Lines* (J. Marshall Unger) – Reviewed by David Cozy
333 *Teaching and Learning in Japan* (Thomas P. Rohlen and Gerald K. Le Tendre, Eds.) – Reviewed by Stephen M. Ryan
337 *The Self-Directed Teacher: Managing the Learning Process* (David Nunan and Clarice Lamb) – Reviewed by Ronald M. Honda
341 *Phonology in English Language Teaching: An International Approach* (Martha C. Pennington) – Reviewed by Ron Grove
344 *Vocabulary, Semantics, and Language Education* (Evelyn Hatch and Cheryl Brown) – Reviewed by Mark O'Neil
347 *Disorders of Discourse* (Ruth Wodak) – Reviewed by Sandra Ishikawa
352 Books to Review

JALT Journal Information

354 Information for Contributors

All materials in this publication are copyright ©1997 by their respective authors.

Japan Association for Language Teaching

JALT is a professional organization dedicated to the improvement of language learning and teaching in Japan, a vehicle for the exchange of new ideas and techniques, and a means of keeping abreast of new developments in a rapidly changing field. Formed in 1976, JALT has an international membership of more than 3800. There are currently 37 JALT chapters throughout Japan. It is the Japan affiliate of International TESOL (Teachers of English to Speakers of Other Languages) and a branch of IATEFL (International Association of Teachers of English as a Foreign Language).

JALT publishes *JALT Journal*, *The Language Teacher* (a monthly magazine of articles and announcements on professional concerns), *JALT Applied Materials* (a monograph series), and JALT International Conference proceedings.

The JALT International Conference on Language Teacher/Learning and Educational Materials Exposition attracts some 2000 participants annually. Local meetings are held by each JALT chapter, and National Special Interest Groups disseminate information on specific concerns. JALT also sponsors specials events, such as conferences on specific themes.

JALT provides awards for Research Grants and Development, announced annually at the conference.

Membership, open to those interested in language education, includes enrollment in the nearest chapter, copies of JALT publications, and reduced admission to JALT sponsored events. For information on membership, contact the JALT Central Office.

JALT National Officers (1997)

President: Gene van Troyer

Vice President: Brendan Lyons

Treasurer: Larry J. Cisar

Recording Secretary: Jim Chambers

Programs: Carolyn Latham

Membership: Richard Marshall

Public Relations: Jane Hoelker

Chapters

Chiba, Fukui, Fukuoka, Gunma, Hamamatsu, Himeji, Hiroshima, Hokkaido, Ibaraki, Iwate, Kagawa, Kagoshima, Kanazawa, Kitakyushu, Kobe, Kochi, Kyoto, Matsuyama, Miyazaki, Nagano, Nagasaki, Nagoya, Nara, Niigata, Okayama, Okinawa, Omiya, Osaka, Sendai, Shizuoka, Suwa, Tochigi, Tokushima, Tokyo, Toyohashi, West Tokyo, Yamagata, Yamaguchi, Yokohama

National Special Interest Groups

Bilingualism; Computer Assisted Language Learning; College and University Educators; Global Issues in Language Education; Japanese as a Second Language; Junior and Senior High School Teaching; Learner Development; Material Writers; Other Language Educators; Professionalism, Administration and Leadership in Education; Second Language Literacy; Teacher Education; Teaching Children; Testing and Evaluation; Video

JALT Central Office

Junko Fujio, Office Manager
Urban Edge Building, 5th Floor
1-37-9 Taito, Taito-ku, Tokyo 110 Japan

JALT Journal

Volume 19, No. 2
November, 1997

Editor
Tamara Swenson
Osaka Jogakuin Junior College

Associate Editor
Sandra Fotos
Senshu University

Assistant Editor
Nicholas O. Jungheim
Aoyama Gakuin University

Reviews Editor
Thomas Hardy
Tamagawa University

Japanese-Language Editor
Naoko Aoki
Shizuoka University

Editorial Board

Charles E. Adamson, Jr.
Miyagi University
Fred Anderson
Fukuoka University of Education
William Bradley
Ryukoku University
James Dean Brown
University of Hawaii, Manoa
Christine Pearson Casanave
Keio University, SFC
Ann Ediger
Teachers College, Columbia University
Rod Ellis
Temple University
John Fanselow
Teachers College, Columbia University
Kevin Gregg
Momoyama University
Dale T. Griffee
Segakuin University
Mary Goebel Noguchi
Ritsumeikan University
David Nunan
City University of Hong Kong
Akito Ozaki
Nagoya University
Tim Riney
International Christian University
Carol Rinnert
Hiroshima City University
Thomas Robb
Kyoto Sangyo University
Patrick R. Rosenkjar
Temple University Japan
Tadashi Sakamoto
Nanzan University
Bernard Susser
Doshisha Women's Junior College
Satoshi Toki
Osaka University
Deryn Verity
Warsaw Medical Center for Postgraduate Studies

Additional Readers: *William Acton, Steve Cornwell, Michael Higgins, Nöel Houck, Peter Robinson, Brad Visgatis*

***JALT Journal* Proofreading**: Greta Gorsuch, Takuo Kinugawa, Brad Visgatis, Jack Yohay

JALT Publications Board Chair: William Acton

The Language Teacher **Editor**: Steve McGuire

Layout: The Word Works **Cover**: Amy Johnson **Printing**: Koshinsha

In this issue

Articles

Five articles are included in this issue. **Dale T. Griffee** examines the continued publication of research based on questionnaires which do not report reliability or validation. He then examines the process of creating, revising, and validating a questionnaire, especially establishing content validity, criterion validity, and construct validity, and provides the steps for teacher-researchers to follow in constructing valid and reliable research questionnaires.

Ahmad Abu-Akel first examines research for insights into the relationship between reading and writing, characterizing ways to conceptionalize this relationship. He then reports the results of a study on the reading-writing relationship for Arabic and Hebrew native speakers studying English in Israel.

Popular songs use in the L2 classroom is examined by **Kim Kanel**, who reports a study comparing the progress in listening comprehension for two groups: one given listening practice with conventional (non-musical) materials, and the other given listening practice with popular song gap-fill exercises. Results indicate that both groups improved equally on a standard measure of listening ability.

The benefits of using reflective journals in a tertiary education environment are examined by **Harumi Moore**. In her study, she found that the use of the reflective journal enhanced learner consciousness in cognitive and metacognitive learning, aided teachers in identifying and analyzing learning issues, and fostered a cooperative relationship between students and teachers as co-participants in the learning process.

One article in Japanese is included in this issue. **Fusako Osho**, **Hiromi Masumi-So**, and **Chihiro Kinoshita Thomson** discuss the design and delivery of a course in "Hospitality Japanese." The researchers first conducted a needs assessment, prepared instructional objectives covering three areas of linguistics, and organized and delivered the course using local resources. However, the evaluation of the course showed improvement was needed in selection of teaching materials and in the balance between the linguistic and non-linguistic objectives.

Perspectives

Three articles are included. First, **John B. Kemp**, examining the importance of empathy in a cross-cultural setting, suggests the use of balanced and informed empathy can often facilitate the resolution of classroom communication breakdowns and can make a significant contribution to ongoing teacher development. Then, **Timothy J. Korst** examines the issue of a silent response to teacher questions in the Japanese

EFL university classroom and presents three activities designed to introduce sociolinguistic skills into the communicative syllabus. Finally, **John Esposito** contrasts the rhetorical traditions of Japanese and English, suggests that during the pre-writing stage of paragraph composition haiku can be used as a complement to or substitute for brainstorming, and provides a sample lesson exploring the practical applications of this approach.

Reviews

Texts on course development, regulation of language use, literary metaphor, functional analysis, business English, literacy reform in occupied Japan, teaching in Japan, managing the classroom, teaching phonology, exploring sematic and lexical theory, and analyzing discourse are reviewed by **Greta J. Gorsuch**, **Virginia LoCastro**, **Valerie Fox**, **Wendy L. Bowcher**, **Steffen Eckart**, **David Cozy**, **Stephen M. Ryan**, **Ronald M. Honda**, **Ron Grove**, **Mark O'Neil**, and **Sandra Ishikawa**.

From the Editors

Correction

The following appendix was inadvertently omitted during the production stage of *JALT Journal, 19, volume 1*. It is included here. We apologize for any inconvenience this omission may have caused.

Influence of Learning Context on Learner's Use of Communication Strategies

Ryu Kitajima

Included in *JALT Journal, 19*(1), 7-23.

Appendix: Examples of Communicative Activities

1. Problem-solving/Debates

The class watched a video presenting evidence by both parties in a lawsuit on an American made-for-TV court case. Since this type of extemporaneous speech presented challenges, a comprehension check was conducted with a series of question-answers between the instructor and the class, followed by a summarization of the content by the class. Then, the class was divided into the plaintiff's side and the defendant's side and debated the case.

In addition, the class debated various issues such as "Advantages and disadvantages of a rural versus an urban life," "Is restriction of TV viewing necessary?" "Which one would you choose, traveling to Hawaii or to Florida?" "Which one is healthier, Italian food or Japanese food?" and "Which one would you choose, living alone or living with a friend?" Materials dealing with these issues (uncaptioned cartoons, short reading passages, audio-tapes) were introduced prior to the debates.

2. Picture descriptions

Pictures were used to provide the students with opportunities to negotiate meaning with each other. In one activity, each subject in a pair had a picture which was identical to her partner's except for 10-15 details. While looking at and describing their own pictures, the pair attempted to identify as many of the differences as possible within a specified time limit. In another activity, one student in a pair had a picture, while the other had a piece of blank paper. According to the information provided by the one with the picture, the other attempted to draw it as accurately as possible.

3. Cross-word puzzles

The students were placed in pairs or groups of three. Each received a slip of paper listing concrete and abstract words and names of well-known people. In turns, students defined a word or described a person on their list while their partner(s) listened in order to determine what was defined or who was described.

4. Problem-solving tasks

The students in the groups discussed potential problems such as a food shortage (during a six month yacht trip around the world) and an accident (while climbing the Himalayas) as a consequence of actions as well as their solutions.

Articles

Validating a Questionnaire on Confidence in Speaking English as a Foreign Language

Dale T. Griffee
Seigakuin University

Despite repeated calls for reliability and validation of data elicitation instruments, research continues to be published based on questionnaires which do not report reliability or validation. The purpose of this paper is to examine the process by which a questionnaire, in this case one designed to measure confidence in speaking English as a foreign language (CSEFL), can be created, revised, and validated. Special attention is given to content validity, criterion validity, and construct validity. The concept and definition of validity is discussed and specific steps and procedures for the validation process are given. A pilot study is briefly summarized followed by the results of the present study. It is concluded that while the majority of the questionnaires used in ESL classroom research in Japan are not valid, the present study provides the necessary steps and procedures by which teacher-researchers can construct valid and reliable research questionnaires.

データ収集の道具の信頼性と妥当性の必要性が繰り返し叫ばれるにもかかわらず、信頼性と妥当性に関する言及がなされていないアンケートの結果にもとづいた研究が出版されつづきている。この論文の目的は、外国語として英語を話す自信（CSEFL)を測定するアンケートが作成され、変更され、妥当性をもつようになる過程を検討する。特に、内容的妥当性、基準関連妥当性、構成概念妥当性に焦点があてられる。妥当性の概念と定義が論じられ、妥当性を達成するための具体的な段取りが提案される。初めにパイロット・スタディが簡単に言及され、その後、本研究の結果が示される。結論としては、日本における英語の授業研究において使われているアンケートの大半は妥当ではないが、ここで示される段取りに従うことで、調査・研究を行なう教師は、妥当で信頼性のあるアンケートを作成することができるということである。

BEST COPY AVAILABLE

JALT Journal, Vol. 19, No. 2, November, 1997

For some time interest in research has been growing among teachers of English as a foreign language (Nunan, 1992, p. xi). As a result of this interest, many classroom teachers have been taking a more active role in conducting and publishing research based on their own classroom observations (van Lier, 1988) and much of this classroom data is being gathered through teacher-designed instruments such as questionnaires and various forms of tests. Many of these instruments, however, are reported with little or no mention of either validity or reliability, which weakens any research based on them (Benson, 1991; Greer, 1996; Keim, Furuya, Doye & Carslon, 1996; Kobayashi, 1991; Teweles, 1996).

First I will begin by discussing the concepts and definitions of validity and reliability, next I will describe the steps and procedures involved in validating a questionnaire, and finally I will report a study aimed at creating a valid and reliable questionnaire. My major purpose is to argue for the role of validity and reliability tests in creating and reporting questionnaire research.

Definitions of Validity and Reliability

In validation, we are interested in arguments which show the degree to which an instrument measures what its author claims it to be measuring (Hatch & Lazaraton, 1991, p. 37; Most & Zeidner, 1995, p. 493). Although it is common to talk about instrument validation, validity is not a quality that belongs in some special way to an instrument. We cannot say that an instrument itself is valid or invalid, but rather that the instrument scores are valid for certain purposes (Cronbach, 1990, p. 145). For example, a proficiency test such as the TOEFL might be considered valid for approximating English proficiency but not for indicating ability to adapt to and live in an English speaking culture. In this sense, validity refers not to the instrument, nor to the scores, but to the use of the scores. More specifically, validity refers to inferences one makes using the scores of a certain test (Pedhazur & Schmelkin, 1991, p. 80).

The notion that there are different types of validity is controversial. Some researchers (Hatch & Lazaraton, 1991) state that there are different types of validity while others (Bachman, 1990; Most & Zeidner, 1995; Pedhazur & Schmelkin, 1991) claim that the notion of different types of validity is mistaken. Either way, it is thought important to report more than one type of validation process. As Bachman says, "it is only through the collection and interpretation of all relevant types of information that validity can be demonstrated" (1990, p. 237).

Important aspects of validation are content validation, criterion validation, and construct validation. Content validity has to do with how well an instrument measures what it says it is measuring (Brown, 1988, p. 102). Brown says that the first step is to establish what the instrument is measuring and the second step is to gather a panel of experts to judge the match between the individual items and the subsections of the instrument. To the extent the panel agrees, one can claim content validity. Criterion-related validity has to do with the extent to which a relationship exists between a high or low score on an instrument and an external criterion believed to indicate the ability being tested or measured. The most common type of criterion-related validation is to compare a new instrument against an established, reliable, and validated instrument. The problem is finding a criterion that is generally accepted and therefore valid because, as Kline (1995, p. 512) states, "the vast majority of psychological tests are not valid." Construct validity, considered central to the validation process (Bachman, 1990, p. 254), is the degree to which the instrument measures the construct under consideration. Construct validation is demonstrated through an argument that the construct, which we cannot see or measure, is indirectly being measured by questionnaire items, which can be seen and measured.

Reliability, on the other hand, is a statistical procedure that indicates how dependably an instrument measures what it claims to be measuring (Brown, 1988; Griffee, 1996a, 1996b; Hatch & Lazaraton, 1991). For any research instrument, including those created by teachers for the purpose of classroom data gathering, one should report both validity and reliability (this, of course, does not include questionnaire forms used only for pedagogical purposes within the classroom). Without such reporting, the reader cannot know how to interpret the inferences made on the basis of the data (Bachman, 1990, p. 24). To put it more bluntly, reliability is a necessary, but not sufficient precondition for validity. If a questionnaire is not reliable, it cannot be valid (Pedhazur & Schmelkin, 1991, p. 81).

It is not the responsibility of the reader to assume reliability (or validity, for that matter); both must be reported. There have been repeated calls for reporting of both validity and reliability (Chaudron, 1988; Kasper & Dahl, 1991; Long, 1990; Luppescu & Day, 1990); these calls apparently are not having much effect among classroom practitioners as evidenced by a check of *The Language Teacher*, a monthly classroom teacher journal published in Japan. From 1976 to 1996, not one of the 13 articles employing questionnaire data in their findings reported instrument reliability or offered any evidence of validation. For the same period, of the

12 articles *JALT Journal* which used data from questionnaires, none reported reliability and nine made no mention of validity. In considering how to construct a questionnaire instrument for research, the literature below suggests five stages of development: the before-writing stage; the writing stage; the piloting stage; the reliability determination stage; and the validation stage.

The Before-Writing Stage—Psychological Constructs

To understand validation, it is necessary to understand what a psychological construct is. A psychological construct is "a theoretically existing (but unobservable) variable" whose existence can be inferred from a variety of sources (Slavin, 1992, p. 244). In the language teaching profession, teachers commonly discuss such psychological constructs as intelligence, aptitude, motivation, confidence, and proficiency. Questionnaires ask specific questions in an attempt to measure such constructs.

Recall that validity is the degree to which inferences can be made about what an instrument claims to be measuring (Ary, Jacobs, & Razavieh, 1990, p. 256; Brown, 1996, p. 231). While validity is not proof, it is an argument on the basis of which researchers hope to convince their readers that the instrument is being used in the situation for which it was designed. In the case of a new instrument, validity is demonstrated through an argument that the instrument is correctly designed for the purposes the researcher has in mind. In order to argue that an instrument is measuring what the researcher states it is measuring, the researcher must make clear what construct is being measured by the instrument. It is for this reason that Bachman (1990) suggests that a first step in instrument creation is to examine theories that discuss what we intend to measure. If no relevant theory exists, Bachman suggests that we could at least create a definition of what we are trying to measure and list the content areas. These content areas can then become the subtests of our instrument (Brown, 1988). For example, suppose that a researcher wants to measure the construct "confidence." He or she examines the theoretical literature on the subject and perhaps finds a paper that defines the term and argues that confidence is composed of qualities X and Y. It is not possible for researchers to directly examine or measure the construct of confidence in students. Nor is it possible to directly measure qualities X and Y. But qualities X and Y are more specific than the construct, and items can be devised that infer the existence of quality X and quality Y. In this way, X and Y have become the basis for the subsections of the instrument. The instrument will have two sections, a section with items purporting to measure quality X and another section composed of items purporting to measure quality Y.

In addition to a serious consideration of the construct, it is also necessary to think about such issues as the requirements for classroom use. For example, how many pages will the instrument contain? Will negative questions be allowed? And what is the type of data desired? (e.g., Likert scales, cloze passages, or open-ended questions) (Tullock-Rhody & Alexander, 1980). In thinking about items which might be included in a questionnaire, Allen (1995) suggests brainstorming items from researcher intuition as well as gathering items from the literature. Another way to elicit items is to ask students similar to those for whom the questionnaire is being developed for items (Horwitz, 1988). For example, in describing a reading questionnaire designed to distinguish good readers from poor readers, Tullock-Rhody and Alexander (1980) report sessions in which they asked elementary school children to describe someone they knew who was a good reader and someone they knew who was a poor reader. Students' views were incorporated into their questionnaire using the students' own language as much as possible.

The Writing Stage

Brown (1996, p. 233) suggests arranging the content areas previously identified and deciding how many items would be needed in each category. Brown also suggests asking colleagues to help in writing items and writing one-third more items than deemed necessary. If some items are not adequate, they can be eliminated. Logically analyze your scoring procedures (Pedhazur & Schmelkin, 1991). Can your construct be measured by "yes" or "no" questions or do you require a greater range of possible responses? If you select a Likert scale, ask a knowledgeable colleague if your scale logically covers all responsible responses in an equal fashion. Try to avoid conflating categories in your instructions to respondents. An example of a conflating or confusing category would be asking respondents if they "believe and approve of" certain practices because it is possible to believe X without approving of X. For example, it is possible to believe that persons should be allowed to smoke cigarettes without approving of smoking. After items have been written, ask expert judges, persons who might be expected to be interested in and experienced with the construct your instrument is attempting to measure, to evaluate your items against the construct. In our imaginary example above, expert judges would be asked to evaluate each item in the subsection against the quality that subsection is attempting to measure. The issue could be stated, do the items in the X section actually measure quality X? If a number of judges object to a given item, serious consideration should be given to either revising or eliminating the item. When all items have been vetted, show them to students similar to the ones for whom the instrument is designed. Ask these students to check each item for comprehensibility and to indicate any vocabulary item they do not understand. It may be necessary to substitute easier vocabulary items or to paraphrase certain items, but a higher level of understanding on the part of respondents will result in less guessing, which in turn will result in higher instrument reliability.

The Piloting Stage

Pilot the instrument on the same type of students for whom the instrument is being designed. In the pilot study, consider writing similar items, placing them in random order, and then correlating student answers to these paired items to see if students answered them in the same way (Reid, 1990; see also Griffee, 1996a). A high correlation between paired items indicates that students interpret the items in a similar way. A low or negative correlation indicates that students are not answering the items in a similar way, which becomes a source of randomness or unreliability. As an alternative, you can correlate each item with the total test scores and keep only the items with high correlations (Cronbach, 1990, p. 170). Revising or eliminating items having low correlation will tend to have the effect of making questionnaire items more consistent and thus more reliable.

The Reliability Stage

With the results of the pilot study, calculate descriptive statistics, reliability coefficients, and the standard error of measurement (Brown, 1996; Griffee, 1996a). What constitutes an adequate reliability coefficient depends on at least six factors: the type of decision, the importance of the decision, the type of reliability estimate, the construct being measured, the instrument medium, and the amount of error the researcher is willing to accept (Griffee, 1996b). The type of decision refers to whether the instrument is being used to measure individuals or to compare groups. Making decisions about individuals demands higher reliability than comparing groups (Pedhazur & Schmelkin, 1991, p. 109). Importance of decision refers to how serious the decision is and how irrevocable the decision is (e.g., acceptance into or rejection from a program). Serious, irrevocable decisions demand higher reliability because of the effect of the decision on individual lives. The type of reliability refers to the formula being used to calculate the coefficient or to the type of reliability calculation (e.g. test-retest, internal consistency). For example, the Kuder-Richardson 21 formula tends to underestimate reliability compared with the Kuder-Richardson 20 formula. The construct being measured refers to whether the construct is easy to measure or difficult to measure (e.g. a mood, feeling, or trait). We may tolerate lower reliability for a difficult-to-measure construct than we will for an easy-to-measure construct. The instrument medium refers to whether the instrument is paper-and-pencil or an interview. An interview might be allowed lower reliability than a paper-and-pencil test. Finally, a researcher may accept lower reliability in an early phase of the research than at a later phase. Table 1 summarizes these comments. There is no hard and fast rule on what constitutes acceptable reliability. Although some writers (Vierra & Pollock, 1992, p. 62) suggest .70 as a cutoff point, others (Ary, Jacobs, & Razavieh, 1990, p. 282) would allow lower levels of reliability, from .30 to .50, for decisions about groups. Finally, Pedhazur & Schmelkin (1991, p. 104) discuss various formulas for determining the reliability coefficient and conclude that Cronbach's alpha is the coefficient of choice when measuring constructs.

Table 1: Factors to consider in determining adequate reliability

Factors	Operationalized as	Reliability could be	
		higher	lower
1. The type of decision	Who/what being measured?	Individual	Group
2. The importance of the decision	What is being decided?	Serious	Not serious
3. The type of reliability	Which formula is used?	KR-20	KR-estimate
4. The construct being measured	Is it difficult or easy to measure?	Easy	Difficult
5. The instrument media	Paper & pencil or interview?	Written	Interview
6. The amount of error the researcher is willing to accept	What stage is the research at?	Late	Early

Table 2: Steps in creating a valid and reliable questionnaire

Stages and procedures

Before writing

1. Investigate available theories that describe your construct.
2. Review all instruments purporting to measure your construct.
3. Define the construct you are trying to measure.
4. List classroom requirements and type of data you want.
5. Brainstorm items from self and literature.
6. Interview colleagues and students for items.

Item writing

7. Decide how many items are required for each subtest or content area, then write more items than are needed.
8. Ask your colleagues for help in item writing.
9. Logically analyze the scoring procedures.
10. Ask expert judges and students to review items.

Piloting

11. Consider pairing and correlating items. Correlate matched pairs, or correlate NS and NNS pairs, or correlate each item with the total, and eliminate or revise low correlating pairs, and pilot again.
12. Pilot the instrument with students similar to those for whom the test is intended.

Reliability determination

13. Calculate descriptive statistics and reliability coefficient.

Validation

14. Explore content validity by convening a panel of experts to judge the match of questionnaire items to construct content.
15. Explore construct validity by conducting a differential group experiment or an intervention experiment.
16. Explore criterion-related validity.

The Validation Stage

It is traditional to consider three types of validity: content validity, construct validity, and criterion-related validity. Bachman (1990, p. 236) suggests that validation is a unitary concept and argues that all three types of validity must be investigated and reported. Content validity can be explored by convening a panel of experts to judge the degree to which the instrument items actually represent the elements being tested (Ary, Jacobs & Razavieh, 1990; Brown, 1996). Construct validity can be explored by differential group experiments, intervention experiments (Brown, 1996), or factor analysis (Boyal, Stankov, & Cattell, 1995; Kline, 1994). A differential groups experiment compares the performance of two groups on a test, one group which obviously has the construct and another group which obviously does not have the construct. An intervention experiment is similar but uses only one group, for example, first year students at the beginning of the school year and the same students at the end of the school year. If the students score higher with each subsequent instrument administration, a researcher can argue that the construct is being acquired. Construct validity can also be explored by statistical procedures such as factor analysis which seek to locate and identify various factors underlying the construction of an instrument. Criterion-related validity can be explored by demonstrating a relationship between test scores of a pilot group similar to those for whom the instrument is designed and some other criterion instrument which is believed to measure the construct being tested, such as: ability as defined by group membership, a recognized test of the same ability, or success on a task that involves the ability being tested (Bachman, 1990, p. 248).

Table 2 lists and summarizes the general stages and specific steps in creating and validating a questionnaire. While in practice it might not be possible or even desirable to realize all 16 procedures, they are listed here for the sake of completeness.

Pilot Study

A pilot study was conducted (Griffee, 1996c) which formed the background of the present study. Two test sources (Mitchell, 1983; Sweetland & Keyser, 1991) were searched for questionnaires measuring confidence and none were found. It was determined that a questionnaire measuring confidence in speaking English would be constructed. Twenty items were brainstormed and administered to 25 university students. Reliability was calculated using the Cronbach alpha formula and paired items were correlated. A factor analysis was calculated looking for roots greater than one using the oblique transformation method. Three factors were identified with eigen values greater than one suggesting that there are possibly three factors of interest. Two factors were identified as a combination of ability (Factor 1) and willingness to engage with others (Factor 3). Factor two was identified as outgoingness or low anxiety.

The Present Study

The primary purpose of this paper is to explain and demonstrate how a questionnaire can be constructed and validated. The purpose of reporting the present study on the creation and validation of a questionnaire on confidence in speaking English as a foreign language (CSEFL) is to illustrate the steps that were taken. The specific research questions addressed in this study are:

1) What is the degree of content validity of the CSEFL?
2) What is the degree of criterion validity of the CSEFL?
3) What is the degree of construct validity of the CSEFL?

Method

Subjects: There were 250 subjects in this study drawn from four small, private colleges in Saitama, Japan. For the most part, the students were in their first or second year, were in their early 20s, and had a variety of majors. Approximately half of the students were males and approximately half were females. Proficiency scores were not available for all students. The entire sample of convenience consisted of each student in 10 intact classes. See Table 3 for group size, school, and school year.

Materials: Version one of the CSEFL questionnaire from the pilot was taken as the base document. Six items having low correlations were eliminated and a panel of experts which consisted of two English native speaker (ENS) males, two ENS females, two Japanese native speakers (JNS) males,

Table 3: Groups, Schools, School Year of Subjects & Alpha Reliability

Group/College	Number	Year	alpha reliability
1. S. Junior College	20	first	.84
2. M. University	26	second	.88
3. T. I. University	21	first	.90
4. S. University	25	first	.92
5. S. University	16	second	.85
6. T. I. University	25	third	.92
7. T. I. University	39	third	.94
8. S. Junior College	21	first	.70
9. S. University	24	first	.86
10. S. University	33	second	.92

and two JNS females was convened to judge the adequacy of the remaining items. The eight panel members, equally divided by gender and ethnic group to reduce possible bias, were interviewed and as a result, six items were dropped. In addition, one item from the factor analysis did not load on any factor and was cut, leaving nine items from the original questionnaire.

A theoretical model of the construct "confidence" was created which hypothesized three aspects of confidence: ability, assurance, and willing engagement. By ability what is meant a command of grammar, vocabulary, and pronunciation. By assurance what is meant that the speaker has a feeling of security and comfort in speaking English. By willing engagement what is meant the speaker is glad to speak in English with native speakers of English.

To create additional items, five colleagues (one JNS female, two ENS females, one JNS male, and one ENS male) were interviewed asking two questions each: think of a person you know who can speak (English/Japanese) with confidence; what are some specific things they do that make you think they are confident? The JNSs were asked about persons who could speak English confidently and the-ENSs were asked about persons who could speak Japanese confidently. Twenty-four items were gathered from the interviews. In addition, as a class exercise, 16 second-year students were asked the same questions and given time to write their answers. Twenty-three items were collected and combined with the 24 colleague answers and the nine original questionnaire items creating a pool of 56 items. From this pool, 30 items were selected for inclusion in the revised questionnaire: 10 under the ability category, 11 under the assurance category, and nine under the willing engagement category. An additional panel of 12 experts was convened to review the pool of 30 items and make recommendations for exclusion or inclusion in the questionnaire.

Procedures: The questionnaire was given to five teachers at the four schools. After teachers were instructed on the nature and purpose of the questionnaire, they administered the questionnaire in their classes and returned the questionnaire to the researcher, who scored it. To help establish criterion-related validity, teachers were asked to select one or two persons in each class who the teacher believed would score high on the confidence questionnaire and one or two students who would score low. Selection was to occur before the questionnaire was administered.

Analysis: The alpha level was set at .05 and all statistics were calculated using StatView 4.5 statistical program for the Macintosh (Abacus Concepts, 1995). The statistical procedures used were Factor Analysis (FA) and Pearson Product-Moment Correlation. In the FA oblique rotation was used and the factor extraction method was the Iterated Principal Axis method using the squared multiple correlation for estimating the initial commonalties. The number of factors to extract was determined by the number with eigen values greater than one. All data sets were independent and, given the large N size, the assumptions of factor analysis e.g. normal distribution are assumed to have been met.

Results

To investigate the first research question on content validity, a 12-member panel (three ENS women, three ENS men, three JNS women, and three JNS men) was convened. An expert was defined as a person who, because of vocation and professional interest, might reasonably be considered as having both interest and knowledge of the subject area under consideration. The panel was asked to rate all items as to validity on a five-point Likert scale of strongly agree, agree, undecided, disagree, and strongly disagree. Five items received five or more negative votes and were eliminated. From the remaining items, a second CSEFL questionnaire was created with 24 items in the three categories of ability, assurance, and willing engagement, and these were randomly ordered.

The CSEFL questionnaire is designed for typical Japanese university students in Japan. Since this group can be comprised of students from intermediate proficiency to rather low proficiency, it was felt that exposing low-level students to the items would yield useful feedback. Six students (three males and three females) typical of the lower proficiency student who would take the questionnaire were individually asked to read the new 24 item CSEFL and indicate any item or word which was not clear. The students did not reject any item as a whole, but did indicate several specific words which they did not understand. One such vocabulary item was the word "argue" (I can argue in English with native speakers) and another word was "willing" (I am willing to speak to many foreigners). "Argue" was changed to "discuss" and "willing" was changed to "I hope to." After these changes were made another eight students (four males and four females) were interviewed in a similar manner and these eight students did not indicate any difficulty with the revised items.

To investigate the second research question on criterion validity, all teachers were asked to nominate one or two students in each class who they believed would score high on the CSEFL questionnaire, and one or two students they believed would score low. The teachers nominated

Table 4: Teacher Nominations of High & Low Confidence Scorers

Class	*N*	Students Nominated to score high	Actually scored high	Nominated to score low	Actually scored low
1	20	2	1	2	1
2	26	2	2	1	0
3	21	2	2	1	1
4	25	2	2	2	1
5	16	1	1	1	0
6	25	1	1	2	2
7	39	3	3	2	1
8	21	3	1	3	1
9	24	3	1	3	1
10	33	2	1	2	0
Total	250	21	15	19	8
Percent		0.71		0.42	

21 students they believed would score high, and 19 students they believed would score low. Students nominated to score high were judged to have actually scored high if their scores were in the top one-third of the class and those nominated to score low were considered to have actually scored low if their percentage correct was in the bottom one-third of the class scores. Table 4 shows the results. The CSEFL agrees with teacher ratings 71% at the higher end and 42% at the lower end.

To investigate the third research question on content validity, first a Principle Components Analysis (PCA) was used followed by Factor Analysis (FA). Hatch & Lazaraton (1991, p. 493) suggest using oblique rotation

Table 5: Factors and Variance Proportions for the PCA

Factors	Magnitude	Variance Proportion
Factor 1	7.724	.322
Factor 2	2.159	.090
Factor 3	1.178	.049
Factor 4	1.129	.047
Factor 5	1.113	.046

factor analysis (FA) to confirm PCA because FA looks at only common variance and ignores error variance and variance not shared by all the factors. The PCA revealed 12 factors with five factors having eigen values over one. Table 5 shows the five factors, their magnitude, and how much of the total variance they account for.

During data inputting, it appeared that some of the items had been rated by students in a contradictory way. For example, many respondents who consistently circled "undecided," "disagree," and "strongly disagree" for most items circled "agree" or even "strongly agree" for item 15 (At a party, I often talk to someone I don't know in English).

Table 6: Correlations of Each Item with the Total Minus Itself

Item number as it appeared in the original brainstorm list	Item number as it appeared in the questionnaire version 2	Correlation
1	5	.568
2	11	.669
3	17	.458
4	1	.545
5	9	.543
6	21	.486
7	20	.474
8	8	.550
9	19	.467
10	4	.656
11	12	.668
12	13	.573
13	22	.545
14	14	.426
15	6	.499
16	24	.455
17	18	.531
18	15	.033*
19	7	.505
20	10	.544
21	16	.496
22	3	.368
23	23	.435
24	2	.551

Note. * = non-significant correlation, all others significant at $p < .05$.

Table 7: Factors and Variance Proportions for the FA

Factors	Magnitude	Variance Proportion
Factor 1	5.440	.363
Factor 2	1.332	.089
Factor 3	.556	.037

Why would students who consistently indicate that they do not like speaking English suddenly indicate that at parties they would talk to a stranger in English? Perhaps a construct other than confidence is being tapped. Kline (1995) suggests using item analysis to remove bad items and factor the reduced set. Each item was correlated against the total minus itself which resulted in the correlations in Table 6.

Table 6 shows questionnaire items 1-8, which were the items hypothesized to measure factor one (ability), items 9-16, factor two (assurance), and items 17-24, factor three (willing engagement). The five highest correlations in each of the three groups were selected and refactored.

Table 8: Factor Loadings: Oblique Solution Primary Pattern Matrix

	Factor 1	Factor 2	Factor 3
Item 1	.520*	.147	-.036
Item 2	-.804	.731*	-.001
Item 4	.178	.691*	.009
Item 5	.546*	.202	.044
Item 6	.198	.396*	.081
Item 7	.396*	.307*	-.077
Item 8	.714*	-.065	-.006
Item 9	.717*	-.038	-.048
Item 10	.576*	.076	-.001
Item 11	.779*	-.018	.053
Item 12	.043	.379*	.442*
Item 13	.012	.138	.992*
Item 16	-.024	.611*	.104
Item 18	.348*	.214	.080
Item 22	.593*	-.048	.114

Note: * = factor loadings at .30 or higher

FA shows three factors, two of which have eigen values over one. The magnitude and proportion of the variance of all factors can be seen in Table 7. The oblique solution primary pattern matrix, Table 8, shows nine items load on factor one, six items load on factor two, and two items load on factor three.

It was hypothesized that five items would load on each of three factors. Results show that all five of the items predicted to load on ability, did so (items 5, 11, 1, 9, and 8), that three out of five predicted items loaded on assurance (items 4, 12, and 6), but that none of the predicted items loaded on willing engagement. In addition, four items loaded in ways which were not predicted (items 7, 10, 18, and 22). Items 7 and 12 load at significant levels on two factors and were cut from CSEFL version three as well as item 13 which loaded only on factor three. This left 12 items for the working version of the questionnaire which appears in the Appendix as version three.

Discussion

The first question is, what is the degree of content validity of the CSEFL? To bring about content validation, two steps must be taken. One, it must be decided what the instrument is claiming to measure and two, it must be decided how to measure the representativeness of each part of the instrument. Condition one has been met in that a model of confidence in speaking English as a foreign language was created which hypothesized three content areas. Condition two has been met in that a panel of experts rated each item in each of the three content areas. All items in the CSEFL have a high degree of panel approval, thus content validly can be claimed.

The second question is, what is the degree of criterion validity of the CSEFL?

Since there are no known reliable or valid measures of confidence in speaking English as a foreign language, this paper uses teacher judgment as a criterion. Criterion response as measured by teacher judgments of students who will score high and students who will score low was mixed. Teachers were generally able to identify students who would score high, but less able to identify students who would score low. One possible explanation is that the CSEFL questionnaire is valid for identifying speakers who are confident, but is not valid for identifying speakers who are not confident. Another possible explanation is that teachers cannot adequately judge certain types of students who appear as not having confidence when in fact, they do. This researcher marked one female student as being low in confidence whereas her score placed her in about in the middle of the class. In a subsequent class exercise, this student declared herself to be an analytic learner who likes solitary tasks such as reading (Nunan, 1988, p. 91). It may be possible that her learning style was interpreted as lack of confidence. It may be necessary to include learning style in addition to the results of a questionnaire such as the CSEFL in compiling a student profile. Against teacher judgment of high achievement, the CSEFL has a relatively satisfactory rating and thus at least partial criterion-related validity can be claimed.

The third question is, what is the degree of construct validity of the CSEFL? The results of the factor analysis are not as clear as we might wish. The high Cronbach alpha reliability coefficients might indicate that high internal consistency in fact reflects item redundancy in which items are little more than paraphrases of each other (Boyal, Stankov, & Cattell, 1995, p. 436). On the other hand, two of the factors have high predicted loadings, which tends to support the validity of the hypothesized construct. The loadings on the third factor are so low as to indicate not only that is particular factor is not supported but also that no additional factor can be substantiated. It may be the case that the lack of a full theoretical model accounting for and describing the construct of confidence leaves us in ignorance as to additional factors. Finally, Boyal, Stankov, and Cattell (1995, p. 421) indicate that while FA provides evidence as to construct validity, which is important, such evidence alone is insufficient. They maintain that predictive evidence alone is essential, and future research may be necessary along those lines. However, the construct validity, criterion validity, and construct validity obtained in the present study suggest that we can argue for partial construct validation. Taking all three types of validation procedures into consideration, it can be argued that the CSEFL is a valid instrument for purposes of researching groups while maintaining some reservations when it comes to individuals keeping in mind the warning of Bachman and Palmer (1996, p. 22) that "it is important for test developers and users to realize that test validation is an on-going process and that the interpretations we make of test scores can never be considered absolutely valid."

Conclusion

This paper has pointed out that the vast majority of the questionnaires used in ESL and EFL classroom research offer no evidence of validation and that conclusions based on the results of such questionnaires are problematic. There might be at least three reasons for this state of affairs. One is that teacher-researchers do not believe it is necessary to report validity or reliability. Second, validity is seen as residing in the instrument. If the instrument was considered valid in another country for another student population, then it must be valid in this country for our students. Third, and closely related, is the idea that if an instrument has been judged valid once, then it must be valid for all time. None of these assumptions are correct and their combined effect is the continued use of invalid and unreliable instruments which results in flawed research. The present study indicates some of the necessary steps and procedures teacher-researchers can take to promote valid and reliable research instruments.

Acknowledgments

The author wishes to thank J. D. Brown for advice and help on earlier forms of this paper as well as a perceptive, unnamed JALT Journal *reviewer.*

Dale T. Griffee, Associate Professor at Seigakuin University, is author of several ESL textbooks. He is editor of the *JALT Applied Materials* (JAM) series and co-editor, with David Nunan, of *Classroom Teachers and Classroom Research*, JALT, 1997. His major research interests are testing, evaluation, and assessment as well as classroom research.

References

Abacus Concepts, StatView 4.5 (Computer Software). (1995). Berkeley, CA: Abacus Concepts.

Allen, A. (1995). Begging the questionnaire: Instrument effect on readers' responses to a self-report checklist. *Language Testing, 12*(2), 133-156.

Ary, D., Jacobs, L., & Razavieh, A. (1990). *Introduction to research in education* (4th ed.). Harcourt Brace.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford University Press.

Benson, M. (1991). Attitudes and motivation towards English: A survey of Japanese freshmen. *RELC Journal, 22*(1), 34-48.

Boyal, G. J., Stankov, L., & Cattell, R. B. (1995). Measurement and statistical models in the study of personality and intelligence. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 417-446). New York: Plenum Press.

Brown, J. D. (1988). *Understanding research in second language learning*. Cambridge: Cambridge University Press.

Brown, J. D. (1996). *Testing in language programs.* Prentice Hall-Regents.

Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning.* Cambridge: Cambridge University Press.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper Collins.

Greer, D. (1996). Gardner and Lambert in the classroom. *The Language Teacher, 17*(1), 10-14.

Griffee, D. T. (1996a). Reliability and a learner style questionnaire. In G. van Troyer, S. Cornwell, & H. Morikawa (Eds.), *On JALT 95 Curriculum & Evaluation: Proceedings of the JALT 1995 International Conference on Language Teaching/Learning* (pp. 283-292). Tokyo: The Japan Association for Language Teaching.

Griffee, D. T. (1996b). Classroom research, instrument reliability, and instrument revision. *Temple University Japan Working Papers in Applied Linguistics, 9,* 25-41.

Griffee, D. T. (1996c). Validation and classroom research instruments: Process and product. *The Journal of Seigakuin University, 9*(1), 53-70.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics.* New York: Newbury House.

Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal, 72*(3), 283-294.

Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition, 13*(2), 215-247.

Keim, B., Furuya, R., Doye, C., & Carlson, A. (1996). A survey of the attitudes and beliefs about foreign language learning of Japanese university students taking communicative courses. *JACET Bulletin, 27,* 87-106.

Kline, P. (1994). *An easy guide to factor analysis.* London: Rutledge.

Kline, P. (1995). A critical review of the measurement of personality and intelligence. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 505-524). New York: Plenum Press.

Kobayashi, J. (1991). Cross-cultural differences in classroom management: Coping with student silences and communication failures. *The Language Teacher, 15*(6), 17-19.

Long, M. (1990). Second language classroom research and teacher education. In Brumfit, C. & Mitchell, R. (Eds.) *Research in the language classroom* (pp. 161-170). ELT Documents 133. Modern English Publications in association with The British Council.

Luppescu, S., & Day, R. (1990). Examining attitude in teachers and students: The need to evaluate questionnaire data. *Second Language Research, 6*(2), 125-134.

Mitchell, J. V. (Ed.) (1983). *Tests in print III.* Lincoln, NE: University of Nebraska Press.

Most, R., & Zeidner, M. (1995). Constructing personality and intelligence instruments. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of* *personality and intelligence* (pp. 417-446). New York: Plenum Press.

Nunan, D. (1988). *The learner-centered curriculum*. Cambridge: Cambridge University Press.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Pedhazur, E. S., & Schmelkin, L. P. (1991). *Measurement, design, and analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reid, J. (1990). The dirty laundry of ESL survey research. *TESOL Quarterly, 24*(2), 323-338.

Slavin, R. E. (1992). *Research methods in education* (2nd ed.). Boston: Allyn and Bacon.

Sweetland, R. C., & Keyser, D. J. (Eds.). (1991). *A comprehensive reference for assessment in psychology, education, and business* (3rd ed.). Austin, TX: Pro-ed.

Teweles, B. (1996). Motivational differences between Chinese and Japanese learners of English as a foreign language. *JALT Journal, 18*(2), 211-228.

Tullock-Rhody, R., & Alexander, J. E. (1980). A scale for assessing attitudes toward reading in secondary schools. *Journal of Reading, 23*(7), 606-614.

van Lier, L. (1988). *The classroom and the language learner: Ethnography and second-language classroom research*. London: Longman.

Vierra, A., & Pollock, J. (1992). *Reading educational research* (2nd ed.). Scottsdale, AZ: Gorsuch Scarisbrick Publishers.

(Received January 7, 1997)

Appendix: Version 3 of Confidence in Speaking Questionnaire

Confidence in Speaking English v.3

Name ______________________ Student # ________________

How confident are you in speaking English?
Circle your best answer for each statement.

For example:

- I like ice cream.
  Strongly agree   Agree   Undecided   Disagree   Strongly disagree

1. I can be interviewed in English.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
2. I would like to study in an English speaking country.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
3. I like speaking English.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
4. I can discuss in English with native speakers.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
5. When I speak English I feel cheerful.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
6. I can speak English easily.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
7. I can show an English speaking visitor around the campus and answer questions.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
8. I say something to other people in English everyday.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
9. I can give my opinion in English when talking to a native speaker.
   Strongly agree   Agree   Undecided   Disagree   Strongly disagree
10. I look for chances to speak English.
    Strongly agree   Agree   Undecided   Disagree   Strongly disagree
11. I will speak to a group of people in English.
    Strongly agree   Agree   Undecided   Disagree   Strongly disagree
12. I am relaxed when speaking English.
    Strongly agree   Agree   Undecided   Disagree   Strongly disagree

On Reading-Writing Relationships in First and Foreign Languages

Ahmad Abu-Akel
University of California, Los Angeles

Reading and writing are related. Inasmuch as reading and writing are both language processes, we can assume relationships between them. More specifically, since both involve the visual processing of language (as compared, for example, with oral/aural processing of language) we may even postulate certain medium-dependent relationships. However the exact nature of these relationships, as well as the implications of these relationships for teaching methods and materials, remain unclear. Research in the last decade has begun to yield insights into various aspects of the nature of the relationships. This paper first characterizes ways one might conceptualize reading-writing relationships, then discusses general findings from first, second, and foreign language research on the nature of reading-writing relationships, and finally, reports the results of a foreign-language reading-writing relationships study conducted for college Arabic and Hebrew native speakers studying English as a foreign language in Israel.

読みと書きには関係がある。両方とも言語にまつわるプロセスであるという意味で、われわれは、両者に関係があることを前提にできる。具体的にいえば、両方とも言語の視覚的処理を必要とするのだから、何らかのメディア固有の関係があることを疑うこともできる。しかしながら、これらの関係が正確にどんな性格のものであるか、これらの関係が教授法と教材に関してどのような示唆をもつのかは明らかではない。過去10年間の研究によって、この関係の性格のさまざまな側面がわかってきた。この論文は、まず、読みと書きの関係の概念化の方法を記し、次に第一言語、第二言語、外国語の読みと書きの関係について先行研究の研究結果を論じ、最後にイスラエルで英語を学ぶアラビア語またはヘブライ語を第一言語とする大学生を対象に行なわれた外国語による読みと書きの関係に関する研究の結果を報告する。

Many different ways for conceptualizing a relationship between reading and writing exist. For example, one might be primarily interested in writing, and wonder about the correlations of reading to writing, or the influence of the processes and products of reading on writing. That is, one might be interested in reading to write. Or, one might be primarily interested in reading and wonder about the

JALT Journal, Vol. 19, No. 2, November, 1997

correlations of writing to reading, or the influence of the process and products of writing to reading. That is, one might be interested in writing to read. One might assume that reading and writing are two sides of the same coin and focus on their similarities and differences in terms of mental processing, or one might focus on the asymmetric relationship of reading to writing—namely that writers must read, but readers do not necessarily have to write. One might be interested in reading the outcomes or products of writing, and from this perspective, one could be interested in a writer reading his or her own product, and the effects of such reading on revision or subsequent writing. Or, one might be interested in the reading done by others of the written products of writers and of the writing process; in other words, one might be interested in reading as the interpretation of writing. One might be interested in either the cognitive aspects of the relationship between reading-writing as mental processes, or one might be more interested in the social aspects of reading and writing and the role of literacy in culture. Or one might conceptualize the relationship between reading and writing from either a dynamic or static perspective. From a dynamic perspective, one would be interested in how the nature of the relationship changes over time, developmentally, or how it may vary over different situations, purposes, goals, and even over different languages (first or second, or third and fourth languages). Finally, and this list is not meant to be exhaustive but merely suggestive, one might be merely interested in the reading-writing relationship as it applies to what we might characterize as "ordinary" texts (simple narrative or expository texts), or in literary or aesthetic texts. Thus, there are many ways to think about reading-writing relationships, and extant research has indeed taken various orientations to the relationship.

Giving all the different ways of conceptualizing the relationship between reading and writing, one can understand that there are about as many different models of the reading-writing relationship. Each model presents those aspects of the relationship as reflected by the respective conceptualizations. Thus, there is no one model of the reading-writing relationship. Each model presents those aspects of the relationship of specific interest or focus to the researcher who developed it. Every researcher necessarily works within a paradigm, and every model has its own dominant focus.

Regardless of the model(s) to follow, adult learners have two primary sources from which to construct a second language system: knowledge of their first language and input from the second language (Carson, Carrell, Silberst, Kroll & Kuehn, 1990). There is evidence that second learners utilize both of these sources in acquiring second language literacy skills.

According to Stotsky (1984) and Tierney and Leys (1986), reading-writing research in English as a first or native language has shown from correlational evidence that "better writers tend to be better readers (of their own writing as well as of other reading material), that better writers tend to read more than poorer readers" (Stotsky, 1984, p. 16). Krashen similarly reports that "a variety of studies indicate that voluntary pleasure reading contributes to the development of writing ability" (1984, p. 4), and that "several studies report statistically significant correlations between reading ability and writing ability" (1984, p. 5). With respect to experimental studies, extant research suggests that while writing instruction, exercises and practice may improve writing, they may not have significant effects on reading. On the other hand, studies that sought to improve writing by providing reading experiences in place of grammar study or additional writing practice found that these experiences were as beneficial as, or more beneficial than, grammar study or extra writing practice (Weaver, 1994; Zamel, 1992). Thus, while additional writing instruction and practice may improve writing, it may not improve reading. Additional reading, however, improves both reading and writing (Stotsky, 1984; Krashen, 1984). Stotsky concludes "it is possible that reading experience may be as critical a factor in developing writing ability as writing instruction itself" (1984, p. 17).

Several researchers have explored the issues of interlingual and intralingual transfer of literacy skills in the development of second and foreign language proficiency. Interlingual refers to the transfer from L1 to L2 reading, and from L1 writing to L2 writing. Intralingual refers to the transfer within L1 or L2 of reading skills to writing skills and vice versa. Cummins (1981) made a strong case for interlingual transfer of literacy skills. He claimed that there is a cognitive/academic proficiency that is common to all languages and that this common language proficiency allows for the transfer of literacy-related skills across languages. Some empirical studies have supported Cummins' claim (Canale, Frenette, & Belanger, 1988; Edelsky, 1982; Goldman, Reyes, & Verhagen, 1984; Mace-Matluck, Dominguez, Holtzman, & Hoover, 1983). For example, the Mace-Matluck et al. (1983) study examined English literacy among students of Cantonese language background and found a significant correlation between literacy acquired in English and literacy level achieved in Cantonese prior to English instruction. In another study, Hirose and Sasaki (1994) investigated the relationship between Japanese students' English L2 expository writing and L1 writing ability and their L2 proficiency. Their findings were that L1 writing was highly correlated with L2 writing ability and that L2 proficiency contributed to L2 writing quality. However, the transfer of literacy-related skills suggested here is limited by Clarke's (1978) threshold hypothesis (see also Alderson, 1984; Cziko, 1978). McLaughlin's (1987) data also suggest that transfer of literacy skills may not be as automatic as Cummins claims. Thus, the picture of interlingual transfer of literacy-related skills is complicated by the notion of a language proficiency threshold suggested by Cummins (1981), Clarke (1978), and Cziko (1978), and by the possibility that this threshold may be a necessary yet not a sufficient condition for transfer to occur, as McLaughlin (1987) suggested.

Intralingual transfer, that is, the mutual influence of reading and writing in the second language, occurs as a result of literacy events in the second language which provide the learner with information about the forms, function, and processes used in literacy activities in the developing language system. Whatever form this second language literacy input may take, it is almost certainly not the case that second language learners acquire reading skills only from writing. Thus, in addition to whatever interlingual transfer effects there are in the L2 from the L1, there are also intralingual effects within the L2 from the influence of L2 reading upon L2 writing and vice versa.

Sarig (1988), Sarig and Folman (1988), and Folman's (1991a, 1991b) works provide insight into the reading-writing relationships in a second or foreign language. They have investigated several aspects of how academic literacy skills relate to L2. Sarig (1988) presented a case study of writing an L1 (i.e., Hebrew) study-summary for both L1 and L2 (i.e., English) texts as an example of what she called a reading-writing "encounter." Her analysis of mentalistic data protocols with a text processing model showed summarization to be a complex mental process involving a number of "cognitive" moves, and further showed that, in terms of the quality of resulting product, summarization from texts in L1 was closely related to summarization from texts in L2, suggesting a transfer of summarization skills from L1 to L2 reading. Moreover, Sarig and Folman (1988) proposed an Academic Literacy Test (ALT) based on the notion of reading and writing as "one integrative meaning construction process" (1988, p. 2). Folman (1991a) presented empirical evidence not only of the effectiveness of explicit training in academic literacy tasks, but also of the transfer of training in academic literacy, and of specifically explicit training in tackling the ALT tasks, from L1 (Hebrew) to L2 (English).

At any rate, while most reading-writing researchers are immersed in generating a model that encompasses the "enigmatic" relationship between reading and writing, little work is done that deals with reading-writing classrooms. This area of study is of extreme importance especially as we try to integrate not only reading and writing, but all language skills, and language skills across content and curriculum areas–and not have to identify the classroom as either writing or reading. The study reported below will try to shed some light on some of the aspects pertaining to reading-writing classrooms.

The Study

Through examining the first language and foreign language reading and writing abilities of college students studying English as a foreign language, te study attempted to determine the relationships across languages [Arabic or Hebrew (L1) and English (FL)], and across modalities (reading and writing) in the acquisition of English literacy skills on an academic level.

Although some research studies (e.g., Canale, Frenette, & Belanger, 1988; Clarke, 1978; Cziko, 1978; McLaughlin, 1987) have looked at the transfer of literacy skills across languages, and a few studies (Flahive & Bailey, 1988; Janopoulous, 1986) have examined reading-writing relationships in L2, there are virtually no studies that attempted to describe how these two strands are related for the same individual engaged in developing literacy skills in his foreign language. By looking at relationships between reading and writing abilities in both first and foreign language, we can begin to describe the contributions of first language literacy skills and the contributions of foreign language reading and writing experiences to the development of literacy in foreign language. Underlying these issues is the question of the role that language proficiency plays.

In this paper, five basic questions are of interest:

a) Is there a relationship between first and foreign language reading abilities?

b) Is there a relationship between first and foreign language writing abilities?

c) Is there a relationship between reading and writing in the learner's first language?

d) Is there a relationship between reading and writing in the learner's foreign language?

e) Does foreign language proficiency affect interlingual or intralingual transfer?

Method

Subjects: A total of 55 native speakers of Arabic and 45 native speakers of Hebrew participated in the study. All the subjects were second year English students in a teacher training college. All subjects had received formal education in English for at least 10 years; and none was a native speaker of English. The level of education achieved in the first language was nearly equivalent for both groups (high school level). The assessment of the subjects' proficiency in English was based on their grades in the writing course, their grades on the reading course, their grades on the Israeli national English matriculation exam, and a placement test adapted from the English psychometric exam for admissions to universities in Israel. On this basis the students' language proficiency varied from low-intermediate (those who had an average of 50-60 from a maximum of 100) to advanced (over 85). The subjects were each assigned to one of three language proficiency levels: low-intermediate (level 1), with 8 subjects; high-intermediate (level 2), with 61; and advanced (level 3), with 31. Table 1 shows the respective groups according to their FL proficiency and native language.

Table 1: Estimated FL Proficiency
for Arabic and Hebrew Native Speakers

	Arabic Speakers (*n* = 55)		Hebrew Speakers (*n* = 45)	
FL Proficiency	*n*	%	*n*	%
Level 1 (50-60)	4	7	4	9
Level 2 (60-85)	43	78	18	40
Level 3 (85 +)	8	15	23	51

Materials: Materials consisted of writing prompts and cloze passages in both the first and the foreign language. The writing prompts were designed to elicit comparison/contrast rhetorical organization, a common pattern of academic discourse and one that presents a clearly discernible set of tasks. (The comparison/contrast type is organized on the basis of opposing viewpoints, either alternative views giving equal weight to two sides, or a pair of views both clearly favoring one side.) The L1 prompt was administered in Arabic or Hebrew. Subjects were instructed to discuss choices in career selection based on the relative availability of job options. The FL prompt–designed to be addressed in English by an

EFL population–asked subjects to write about the importance of belonging to a group or of being an individual in order to achieve one's goals.

The native language reading passages to be turned into cloze readings were selected by three teachers of each language; three Hebrew teachers and three Arabic teachers. The passages selected followed the following criteria: (a) The topic of the passage must be of general interest; (b) the passage must be authentic text aimed at readers with high school level reading skills; (c) the passage must exhibit comparison and contrast rhetorical organization; and (d) the length of the passage must be between 300 and 400 words. The Arabic article was about differences between Jewish and Arab schools in Israel, and the Hebrew article was about rural versus urban life styles. The English text, selected by English native speakers, discussed the effect of environment on dress codes.

After the passages were selected, the teachers of each language used cloze procedures on the passages, following a 7th word deletion rate and maintaining the first sentence of each passage intact. The English passage contained 52 blanks; the Arabic 44 blanks; and the Hebrew, 44 blanks. Instructions included sample sentences with words written in the blanks. All passages were then typed, and the space allotted for each cloze item was standardized across languages.

Procedures: All writing tasks preceded all reading tasks so the reading passages would not provide models for writing and thereby affect writing performance. L1 and FL tasks were counterbalanced.

Subjects were given between 30 and 45 minutes to complete each of the four tasks. Tasks were administered over a two-week period to ensure that language learning between task administration would not significantly affect results. No dictionaries were allowed, and students were not given additional instructions apart from those appearing with the essay prompts and cloze passages.

Scoring: Both the first language essays (Arabic and Hebrew) and the English essays were evaluated by native speakers of those languages using 6-point scales. Each essay was scored by two raters; essays with scores that differed by two or more points were read by a third rater and the extreme score was dropped. The score for each essay was the average of two raters.

The English essays were scored using a 6-point scale developed to score the Test of Written English (TWE) portion of the TOEFL (Appendix A). All three raters had been trained by the researcher and assistants as to how to score based on the TWE criteria.

Since no guidelines existed for Arabic or Hebrew essay scoring, the raters of these essays developed a scoring guideline by following a two-step process. First they were asked to sort the essays into six piles, with each pile corresponding to a degree of proficiency: Essays ranked 6 were the best, and 1 the worst. Then the raters were asked to write a set of descriptors characterizing the features of each of the six groups of essays, resulting in a written 6-point scale for Arabic (Appendix B) and Hebrew (Appendix C) essays. Though some intrinsic language-specific differences are expected, the three languages' criteria for essay evaluation focusing on coherence, topic development, and language usage were all similar.

The English, Arabic and Hebrew raters were all experienced (minimum five years of experience) in teaching writing classes at the college and university levels. Estimates of interrater reliability (coefficient alpha) for the two primary raters in each essay category are reported in Table 2, along with percent of rater agreement, rater means, and standard deviation. Although a third rater was used to provide as accurate an average holistic rating as possible for use in analyses, coefficient alpha and percent agreement are reported to provide information about the functioning of the three 6-point holistic scoring scales. Rater agreement was operationally defined as ratings within one scale point. Coefficient alphas ranged from .78 (FL essay) to .92 (Arabic essays). The alpha reported for the FL essay raters is low, in part due to the relatively restricted variability of the second rater's ratings (*SD* = .76). The agree-

Table 2: Essay Rater Means, Standard Deviations, Percent Agreement, & Coefficient Alphas

Essay Language	Rater 1	Rater 2	Percent Agreement	Coefficient Alphas
Arabic (n = 55)				
M	3.00	3.2	0.98	0.92
SD	1.6	0.5		
Hebrew (n = 45)				
M	3.1	2.9	0.67	0.87
SD	1.7	1.3		
English (n = 100)				
M	3.5	3.3	0.91	0.78
SD	1.00	0.76		

ment rate between the two FL essay raters of 91% and the reported alpha are evidence of rating system reliability. Either one or both of the raters assigned ratings of 3 to 71% of the FL essays and 4 to 84%.

Because the sample was initially distributed into six categories by raters, the method used for constructing the L1 essay rating scales yielded greater rating variability compared to the FL scales. Coefficient alpha for the Hebrew essays was .87, although the rater agreement was only 67%. The first Hebrew rater consistently rated essays higher than did the second rater. The Arabic essay ratings had both higher interrater reliability (.92) and higher rater agreement (98%). Cloze passages were scored using exact-word scoring, since Oller's (1979) review of cloze research indicated that although percentage scores may be lower with exact-word scoring, rank order should remain the same with exact-word or acceptable substitute scoring.

Results

Mean scores by task are reported in Table 3. The mean for the Hebrew cloze test was 29.9, and for the Arabic the mean was 32.8, out of a total 44 blanks on each test. The differences in means was not significant, as revealed by the *t* test at .05 level of significance, suggesting that the subjects are equally competent in this language skill. The mean score of the Hebrew subjects on the English cloze was 24.5; the Arabic was 21 (52 blanks total), reflecting the different FL language proficiencies of the two groups.

The L1 essay scores (Table 3) were comparable for the two groups: The Hebrew mean was 3.2, the Arabic, 3.3. The English essay scores again reflected the difference in FL language proficiency: the Hebrew subjects' mean rating was 3.6, and the Arabic subjects', 3.1. This difference in FL proficiency between the two subject groups should be kept

Table 3: Task Means & Standard Deviations

Task	Maximum score	Hebrew (*n* = 45)		Arabic (*n* = 55)	
		M	*SD*	*M*	*SD*
L1 cloze	44	29.9	4.1	32.8	3.1
FL cloze	52	24.5	6.1	21	5.2
L1 essay	6	3.2	1.56	3.3	1.5
FL essay	6	3.6	0.75	3.1	0.70

in mind as a possible source of influence on the analysis relating to L1 and FL language skills.

The relationships between L1 and FL reading and writing were investigated initially by examining the correlation coefficients. Weak to moderate correlations are reported in Table 4. Correlations magnitudes for the reading-writing relationship may be considered in terms of Shanahan and Lomax's (1986) proposed model of the reading-writing relationship, which argues for the existence of multiple relations (i.e., interactions among language skills such as word analysis, spelling knowledge, and word recognition may differ within and across discourse levels), as well as the possibility that the nature of the reading-writing relationship might change with development and thus not be linearly related. In this case, the Pearson correlation thus may underestimate the actual relationship between these two language skills.

Table 4: Correlations by Language Groups for L1 and FL Reading and Writing Tasks

	Hebrew ($n = 45$)	Arabic ($n = 55$)
L1 reading x FL reading	$r = 0.37^{*}$	$r = 0.51^{**}$
L1 writing x FL writing	$r = 0.02$	$r = 0.23^{*}$
L1 reading x L1 writing	$r = 0.30^{*}$	$r = 0.50^{**}$
FL reading x FL writing	$r = 0.54^{**}$	$r = 0.27^{*}$

$^{*}p < .05$
$^{**}p < .01$

The correlations in Table 4 show the following relationships: (1) L1 and FL reading scores had weak to moderate correlations for both the Hebrew and Arabic subjects; (2) L1 and FL writing scores showed weak positive correlations for the Arabic but not for the Hebrew subjects; (3) L1 reading and writing showed weak to moderate correlations for both groups, as did FL reading and writing. For both groups, there are stronger relationships between reading abilities across languages than between writing abilities across languages. The L1-FL writing relationships for Arabic is weak, and for the Hebrew subjects it is not significant. The correlations in Table 4 also show that for the Hebrew subjects the relationship between reading and writing is strongest in FL, but for the Arabic subjects the reading-writing relationship is strongest in L1.

Reading and writing are related, but the strength and nature of the relationship differs for each of these groups, either due to language or other background variable differences. In this respect, one should probably mention that the difference between the Arabic and Hebrew groups could lie in the fact that though English, is by definition, a foreign language for both groups, for the Arab students English is taught as a third language after the Arabic and Hebrew languages, and for the Hebrew subjects it is only the "second language," after Hebrew. Or it could be attributed to some idiosyncratic writing styles different languages have. That is, superb writing in Arabic is dependent on the use of highly elaborative and descriptive vocabulary. Moreover, Arabic writing is not direct, and is rather manipulative. The ability to manipulate language is measured against writing quality. Thus, and as can be noticed from the evaluative criteria on the use of vocabulary (Appendices A-C), good English writing resembles to some extent that of Hebrew, and both are different from Arabic. Therefore, the differences in results between Arabic and Hebrew subjects may in part be attributed to these sets of circumstances.

The means and correlations by FL proficiency levels showed pattern differences by language groups and by levels. However, because of the *n*-sizes for Level 1 (4 each for Hebrew and Arabic) and for the level 3 Arabic group (8), it is difficult to draw conclusions about the role that language proficiency plays in these reading-writing relationships. At any rate, two trends, although not statistically significant, did appear that are worth noting. First, FL reading and writing scores tended to increase as FL proficiency increased. This trend is confirmed by Hirose and Sasaki (1994) who report that Japanese EFL students' general L2 proficiency contributed significantly to the quality of the their L2 writing. Similar results were also reported by Cumming (1989) and Pennington and So (1993).

Second, however, L1 reading and writing scores tended to decrease as FL proficiency increased. This was particularly noticeable for L1 writing, where means for both groups decreased from an average of 3.4 at level 2 proficiency, to 2.8 at level 3. It appears that L1 writing skills are rated weaker as L2 proficiency increases. Potentially, there could be a number of explanations for this trend. It is possible that this is a reflection of the fact that students in an FL academic environment (obtaining a degree in the English language) are not engaged in L1 academic writing activities of the type we are measuring (comparison/contrast). In fact, Abu-Akel (1996) has reported a correlation between the rhetorical organization of the text and one's reading and/or writing ability. The resulting attrition may be similar, then, to the phenomenon of language loss that occurs when language is no longer used sufficiently to maintain proficiency. This seems particularly true for the Arabic subjects, whose writing and speaking modes are completely different (i.e. diglossia). Still, these results by language proficiency must be interpreted cautiously, given the low numbers of subjects on these levels.

Discussion

The data suggest that interlingual transfer can occur, but that the pattern and the strength of this pattern varies according to first language background and other aspects of educational background. For reading, the transfer from L1 to FL was similar for both Hebrew and Arabic subjects, but for writing, the transfer from L1 to FL was different. These differences may be a function of FL language proficiency. Another possibility, though it is not investigated here, is that cultural differences are reflected in the literacy practices and abilities of the two groups. There is more "cultural overlap" between Hebrew and English than for Arabic and English (Abu-Rabia, 1995). Abu-Rabia (1995) has found that cultural background and social contexts contribute either negatively or positively to L2 learning: the greater the "cultural overlap" the more positive the contribution to one's L2 learning. Altarriba and Forsythe (1993) also contend that cultural schemata has bearing on one's ability to read and write in L2. Lack of knowledge of cultural schemata may obscure one's understanding of the writer's message, or result in an inability to express oneself in a manner that is appropriate for that culture. In more general terms, anybody who has tried to learn a second language to any considerable depth, particularly where there is little "cultural overlap," say English-Arabic or English-Japanese rather than Italian-French, will recognize that learning and using idioms, for example, involves attaining a deep understanding of the social practices which underlie the use of any particular expression in a specific context. More insight into this area indeed calls for further research.

Beyond cultural and proficiency variables, some of the differences between Arabic and Hebrew students could be related to the functional nature of these languages. Arabic is a classic diglossic language, where the spoken mode (the Low variety) is not written, and Classic Arabic (the High variety), is learned as a second language and used as the writing mode (Ferguson, 1991). In a study conducted by Abu-Asbi (1995), Arabic diglossia emerged as a significant factor effecting Arabic speakers' proficiency in English. An Arabic student not only has to transfer listening and speaking skills from first to second language, as is the case for the Hebrew speaker, but in fact as a reader/writer has to transfer second language skills to a third language. In this respect, and as argued by Geva and Ryan (1993) the number of languages learned could interfere with one's extent of proficiency in any one language.

The results also suggest that reading ability transfers more easily from L1 to FL than does writing ability. In fact, a weak relationship for L1-FL writing is indicated by the correlations for both groups. The results pertaining to the contribution of L1 writing to L2 writing in this study fail to confirm the results reported by Hirose and Sasaki (1994). They report that L1 writing ability significantly contributes to L2 writing ability. The difference between the present study and theirs remains obscure for there could be a variability in the definition employed for whether or not language proficiency evaluations are comparable in both studies.

At any rate, further research is needed to determine whether the different variables that predict Hebrew and Arabic writing scores are the result of FL proficiency, cultural differences, or the diglossic situation in writing skills. It seems that L1 and FL educational levels interact in various complex ways.

Conclusions and Pedagogical Implications

Given the exploratory nature of this study, any teaching implications based on these preliminary findings should be treated with caution. Still, the results suggest some general implications for the classroom.

As other studies have suggested (e.g., Carson et al., 1990), there are significant correlations between L1 and FL reading for both Arabic and Hebrew groups. That is, there is a positive relationship between reading in the first and reading in the foreign language. Although other factors may be important, the relationship could and should be exploited in FL reading pedagogy. L1 reading skills can and should be used in FL reading pedagogy, but the instructor should not depend on automatic transfer of L1 reading abilities/skills to FL reading. Similar implications are suggested by Carson et al. (1990), however, for ESL Japanese and Chinese adult learners

The weak correlation between L1 and FL writing for the Arabic subjects, and the lack of correlation of L1 and FL writing for the Hebrew subjects (whose proficiency was higher), suggest that the extent to which L1 may be exploited or used in FL writing pedagogy may be limited to lower FL proficiency levels and/or certain L1 language groups. This implication is supported by Kobayashi and Rinnert (1992). In their study, among other things, they investigated the effect of Japanese EFL proficiency on writing quality. Their findings suggest that while lower-proficiency students may benefit from L1 (in the form of translation from L1), higher-proficiency students generally do not benefit very much from it. Hence the writing teacher may rely even less than the reading teacher on the transfer of L1 writing skills to L2 writing.

The differences in the reading-writing relationships between the Arabic and Hebrew groups suggest that if the nature of the L1 and FL reading-writing relationship changes as FL proficiency develops, then the extent to which L1 may be relied on in pedagogy also changes with FL literacy development. That is, whereas teachers may be able to exploit L1 literacy relationships in the transfer of FL literacy practices at lower proficiency levels, they cannot do so reliably at more advanced FL levels. Here, teachers need to rely more on the developing FL literacy. In other words, at lower proficiency levels, interlingual transfer may be more important, where as at higher proficiency levels, intralingual input may be the more significant source for developing FL literacy skills.

Finally, the results reported here can be further enhanced by adding raters and/or adding a wider range of reading and writing topics that would help unravel the effect of different topics on the nature of the reading-writing relationship. Further research should address the issues of L1 and FL scale equivalency and rating variability raised in this exploratory study. Moreover, we need to learn more about the ways in which FL writing skills are affected by interlingual transfer and intralingual input; and to investigate further those literacy practices of these two groups that may relate to different patterns of FL literacy acquisition.

Ahmad Abu-Akel, of Bar-Ilan University, Israel, is currently at the University of California, Los Angeles, Department of TESL/Applied Linguistics.

References

Abu-Akel, A. (1996). The role of schemata in ESL reading comprehension. *The Language Teacher, 20*(9), 6-11.

Abu-Asbi, K. (1995). [The Arabic education system in Israel: From an existing situation towards determining policy and education planning]. Jerusalem: The School for Educational Leadership.

Abu-Rabia, S. (1995). Multi-cultural and problematic social contexts and their contribution to L2 learning. *Language, Culture and Curriculum, 8*(2), 183-198.

Alderson, J.C. (1984). Reading in a foreign language: A reading problem or a language problem? In J.C. Alderson, & A.H. Urquhart (Eds.), *Reading in a* *foreign language* (pp. 1-27). New York: Longman.

Altarriba, J. and Forsythe, W. J. (1993). The role of cultural schemata in reading comprehension. In J. Altarriba (Ed.), *Cognition and culture: A cross cultural approach to psychology* (pp. 145-155). Amsterdam: Elsevier Science Publishers.

Canale, M., Frenette, N., & Belanger, M. (1988). Evaluation of minority student writing in first and second language. In J. Fine (Ed.), *Second language discourse: A textbook of current research* (pp. 147-165). Norwood, NJ: Ablex.

Carson, J.E., Carrell, P.L., Silberst, S., Kroll, B., & Kuehn, P.A. (1990). Reading and writing relationships in 1st and 2nd language. *TESOL Quarterly, 24*, 245-266.

Clarke, M.A. (1978). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning, 29*, 121-150.

Cumming, A.H. (1989). Writing expertise and second language proficiency. *Language Learning, 39*, 81-141.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In The California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3-49). Los Angeles: Evaluation, Dissemination and Assessment Center, California State University.

Cziko, G. (1978). Differences in first and second language reading: The use of syntactic, semantic and discourse constraints. *Canadian Modern Language Review, 34*, 473- 489.

Edelsky, C. (1982). Writing in bilingual program: The relation of L1 and L2 texts. *TESOL Quarterly, 16*, 211-228.

Ferguson, C.A. (1991). Diglossia revisited. *Southeast Journal of Linguistics: Studies in Diglossia, 10*(1), 214-234.

Flahive, D., & Bailey, N. (1988). *Exploring reading/writing relationships in adult second language learning*. Unpublished manuscript.

Folman, S. (1991a). *Attenuated transfer of academic literacy from the L1 to the L2: Some empirical evidence*. [Paper presented at the 1991 TESOL Convention]. New York.

Folman, S. (1991b). Towards an EFL/ESL reading-writing model of academic learning. In J. Carson, & I. Leki (Eds.), *Reading in the composition classroom: Second language perspectives* (pp. 81-93). Rowley, MA: Newbury House.

Geva, E., & Ryan, E. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning, 43*(1), 5-42.

Goldman, S.R. , Reyes, M., & Verhagen, C. K. (1984). Understanding fables in first and second languages. *NABE Journal, 8*(2), 35-66.

Hirose, K., & Sasaki, M. (1994). Explanatory variables for Japanese students' expository writing in English: An exploratory study. *Journal of Second Language Writing, 3*, 203-229.

Janopoulous, M. (1986). The relationship of pleasure reading and second language proficiency. *TESOL Quarterly, 20*(4), 764-768.

Kobayashi, H., & Rinnert, C. (1992). Effect of first language on second language writing: Translation versus direct composition. *Language Learning, 42*(2),183-215.

Krashen, S.D. (1984). *Writing: Research, theory, and applications*. Oxford: Pergamon.

Mace-Matluck, B.J., Dominguez, D., Holtzman, W., & Hoover, W. (1983). *Language and literacy in bilingual instruction*. Austin, TX: Southeast Educational Laboratory.

McLaughlin, B. (1987). Reading in a second language: Studies with adult and child learners. In S.R. Goldman & H.T. Trueba (Eds.), *Becoming literate in English as a second language* (pp. 57-70). Norwood, NJ: Ablex.

Oller, J.W., Jr. (1979). *Language tests at school*. London: Longman.

Pennington, M.C., & So, S. (1993). Comparing writing process and product across two languages: A study of 6 Singaporean university student writers. *Journal of Second Language Writing, 2*, 41-63.

Sarig, G. (1988). *Composing a study-summary: A reading-writing encounter* [Unpublished manuscript].

Sarig, G., & Folman, S. (1988). *Issues in testing academic literacy: An integrative approach* [Unpublished manuscript]. Tel-Aviv.

Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading-writing relationship. *Journal of Educational Psychology, 78*, 116-123.

Stotsky, S. (1984). Research on reading/writing relationships: A synthesis and suggested directions. In J. M. Jensen (Ed.), *Composing and comprehending* (pp. 7-22). Urbana, IL: NCTE.

Tierney, R.J., & Leys, M. (1986). What is the value of connecting reading and writing? In B.T. Petersen (Ed.), *Convergences: Transactions in reading and writing* (pp. 15-29). Urbana, IL: NCTE.

Weaver, C. (1994). *Reading process and practice: From socio-psycholinguistics to whole language*. Portsmouth, NH: Heinemann.

Zamel, V. (1992). Writing one's way into reading. *TESOL Quarterly, 26*(3), 463-485.

(Received January 12, 1997; revised May 10, 1997)

Appendix A: Test of Written English (TWE) Scoring Guidelines*

6 Clearly demonstrates competence in writing on both the rhetorical and syntactic levels, though it may have occasional errors.

A paper in this category
- is well organized and well developed
- effectively addresses the writing task
- uses appropriate details to a support thesis or illustrate ideas
- shows unity, coherence, and progression
- displays consistent facility in the use of language
- demonstrates syntactic variety and appropriate word choice

5 Demonstrates competence in writing on both the rhetorical and syntactic levels, though it will have occasional errors.

A paper in this category
- is generally well organized and well developed, though it may have fewer details than does a level 6 paper
- may address some parts of the task more effectively than others
- shows unity, coherence, and progression
- demonstrates some syntactic variety and range of vocabulary
- displays facility in language, though it may have more errors than does a level 6 paper

4 Demonstrates minimal competence in writing in both the rhetorical and syntactic levels.

A paper in this category
- is adequately organized
- addresses the writing topic adequately but may slight parts of the task
- uses some details to support a thesis or illustrate ideas
- demonstrates adequate but undistinguished or inconsistent facility with syntax and usage
- may contain some serious errors that occasionally obscure meaning

3 Demonstrates some developing competence in writing, but remains flawed on either the rhetorical or syntactic level or both.

A paper in this category shows:
- inadequate organization or development
- failure to support or illustrate generalization with appropriate or sufficient detail
- an accumulation of errors in sentence structure and/or usage
- a noticeably inappropriate choice of words or word forms

2 Suggests incompetence in writing.

A paper in this category is seriously flawed by one or more of the following weaknesses:
- failure to organize or develop
- little or no detail, or irrelevant specifics
- serious and frequent errors in usage or sentence structure
- serious problems with focus

1 Demonstrates incompetence in writing.
A paper in this category will contain serious and persistent writing errors, may be illogical or incoherent, or may reveal the writer's inability to comprehend the question. A paper that is severely underdeveloped also falls into this category.

*Reprinted by permission of Educational Testing Service, the copyright owner

Appendix B: Evaluation Scale Descriptors: Arabic Essay

6 The argument presented is very clear.
The sequencing of words and sentences is consistent and smooth.
The topic is addressed well.
The overall presentation is well organized.
The vocabulary is abundant.
5 The argument is clear.
The persuasion is a little weaker than Level 6.
The fluency of the language is good.
The vocabulary used is not as elaborate as that in Level 6 papers.
4 The overall control of the language is more than the average, but not completely satisfactory.
The argument mostly follows the topic.
The variety and the type of sentence construction used need more consideration.
3 The argumentation, sequencing of the sentences, expression and vocabulary are acceptable.
The level is average.
2 The logical development is missing.
The argument is not clear.
The vocabulary used is limited.
The paper is not fully developed.
1 The topic is not addressed well.
The statements are irrelevant.
The question is misunderstood.
The paper lacks the clear arguments about the topic.

Appendix C: Evaluation Scale Descriptors: Hebrew Essay

6 The essay is well written, characterized by thoughtful and coherent reasoning.
    The essay plan is clearly signaled by transitions.
    The overall presentation of argument is convincing, with varied sentence constructions and persuading evidence.
    The main idea is identified.
    Superior control over language.
5 A clear understanding of the topic is demonstrated.
    The argument is unified and coherent with the subject.
    Opening and closing statements are related to each other.
    Ideas are sufficiently developed.
    There may be some minor errors in usage and sentence structure.
4 The subject is clear.
    Some sequence of ideas.
    The essay gives directions to subsequent reasoning.
    The essay complete the basic task of the assignment.
    Not enough convincing evidence to support the main point.
    Some irrelevant sentences.
3 The subject is identified.
    The main idea is stated.
    Reasoning is not adequate or convincing.
    No exhaustive argument.
2 Little development of ideas.
    The main point is not clear.
    No evidence to support the main idea.
    Some errors in reasoning.
    The topic is limited.
1 Absence of thesis statement.
    The main point is not clearly stated.
    No sequence of ideas.
    No overall presentation of the argument.
    No basic structure of essay.
    Badly mishandled sentence structure.
    Lack of convincing and logic argument.
    The essay is lacking in content.

Teaching with Music: A Comparison of Conventional Listening Exercises with Pop Song Gap-fill Exercises

Kim Kanel
Kinki University

Popular songs in the L2 classroom not only increase interest and motivation, but also serve to meet a number of pedagogical needs. However, for song-based tasks to gain wider acceptance, it must be shown that they are as effective as conventional tasks. This paper reports a study comparing the progress in listening comprehension for two groups: one given listening practice with conventional (nonmusical) materials (n = 358), and the other given listening practice with popular song gap-fill exercises (n = 334). Results on pre- and post-test scores using the Japan Association of College English Teachers (JACET) Basic Listening Comprehension Test indicate that both groups improved equally and made significant progress. On post-treatment questionnaires, the song group expressed both higher approval for the time spent on the tasks and increased interest in English.

第二言語の授業におけるポップ・ソングは、興味と動機を高めるだけでなく、数多くの教育的ニーズにも応える。しかしながら、歌を使ったタスクがより広く受け入れられるためには、伝統的なタスクと同じくらい効果的であることを示す必要がある。この論文は、聞き取りの進歩について二つのグループを比較する。一つは伝統的な（音楽を使わない）聞き取りの練習をさせたグループ（358人）、もう一つはポップ・ソングの穴埋めによって聞き取りの練習をさせたグループ（334人）である。JACETのBasic Listening Comprehension Testによるプレ・テストとポスト・テストの結果は、二つのグループとも同様に、目覚ましい進歩をとげたことを示した。授業後のアンケートでは、歌を使ったグループのほうが、タスクに使った時間をより高く評価し、英語への興味がより増したと答えた。

Research into L2 listening comprehension development has shown that foreign language learners must acquire the ability to: (a) discriminate among the distinctive sounds of the target language; (b) recognize reduced forms of words; (c) distinguish word boundaries; and (d) guess the meanings of words from the contexts in which they occur (see Richards, 1983; Rost, 1990). In most Japanese

JALT Journal, Vol. 19, No. 2, November, 1997

high schools, however, students of English are offered little exposure to the actual sounds of the language in context, and routinely fail to correctly distinguish spoken words they might otherwise understand in written form. In order to provide students with exposure to authentic English, and at the same time stimulate motivation for the lessons, teachers are increasingly utilizing materials based on English language movies and songs. Though both media can provide listening practice, songs have the additional advantage of almost always being heard in the original language (i.e., no L1 subtitles or dubbing). In addition, like movies, the topics and language of pop songs tend to reflect the interests, values, and tastes of young adult EFL learners more accurately than the material used in commercial textbooks (Coe, 1972; Dubin, 1974; Loew, 1979; Murphey, 1988, 1989; Smith, 1976).

In addition to increasing interest in the content of lessons, songs can be used to introduce practically any area of the language learning syllabus (DeSelms, 1983; Dubin, 1974; Sekara, 1985; Urbancic & Vixmuller, 1981). In the last few years several resource books presenting song-based activities for classroom use have been published (Cranmer & Laroy, 1992; Griffee, 1992; Murphey, 1992), in addition to a number of textbooks containing exercises using songs (Berglund, 1983; Dougill, 1989; House & Manning, 1992; Kanel, 1995; 1997; Kanzaki, 1988; Mosdell, 1984; Posener, 1987; Someya & Ferrasci, 1988; Sato & Sasanuma, 1988). There is still, however, a lack of research examining the effects of song use on language acquisition. As with any other teaching method, for song-based activities to gain legitimacy, it must be demonstrated that they are as effective as conventional activities.

The Effect of Music on Cognitive and Affective Variables

The role of music and song in development of human languages, as well as the linguistic development of the individual, has been acknowledged by anthropologists (Murphey, 1990b). Studies of both normal and learning-disabled students suggest that properties of music (rhythm, melody, harmony, dynamics, form and mood) aid the cognitive processing of first language vocabulary, enhance retention and promote overall language development and reading skills (Botari & Evans, 1982; Gfeller, 1983; Isern, 1958; Jalongo & Bromley, 1984; McCarthy, 1985; Schuster & Mouzon, 1982). In addition, music has the potential to break down many of the affective barriers that inhibit learning (Lozanov, 1979; Meyer, 1956; Stoudenmire, 1975), and can make learners more receptive to subject matter by increasing consciousness and emotional involvement in the learning process (Rosenfeld, 1985).

Language Acquisition and Development: The Din and SSIMH Phenomena

Barber (1980) first described a phenomenon which she called the "Din in the head" to account for the involuntary rehearsals that often take place in foreign language learners' minds. Krashen (1983) hypothesized that the Din was a result of the stimulation of the Language Acquisition Device (LAD). Parr and Krashen (1986) later surveyed second language learners, finding that three-quarters have had the Din experience to some degree: frequently occurring after extended periods of comprehensible foreign language input.

Murphey (1984; 1990a) hypothesized that music and song might provide a similar Din, and described what he called the "song-stuck-in-my-head" (SSIMH) phenomenon (i.e., a song or melody we just cannot get out of our heads). In a survey of 30 native speakers of English and 19 native speakers of other languages, he found that all had experienced the SSIMH, and that all but two had had it in a second language. Murphey reasoned that if the Din works to stimulate language acquisition, so then should the SSIMH phenomenon. In contrast to Krashen's concepts of the Din, however, Murphey suggested that the SSIMH phenomenon does not necessarily need comprehensible input. This notion is particularly important when considering students at the beginning levels of foreign language acquisition. Murphey concluded that if prior exposure to language does affect subsequent learning, many EFL/ESL students have already experienced a significant amount of contact with English through songs.

Language Learning Songs vs. Authentic Songs

Because authentic songs sometimes contain non-standard structures and vocabulary, and irregular stress and intonational patterns, some educators have expressed doubts as to whether the potential benefits of their use outweigh the possible detrimental effect of an incorrect model of the spoken language (Coe, 1972; Jolly, 1975; Richards, 1969). Richards (1969) advised that control should be applied to the language of songs, just as to any other part of the English course, and called for the creation of special English language learning songs. During the 1970s and early 1980s, numerous ESL song books were published that did contain controlled language and few, if any, of the "mistakes" found in authentic songs.[1]

A number of educators, however, began to counter previous objections to authentic songs, basing their arguments on the changing premises of post-Audiolingual Method approaches to language learning. Dubin (1974) writes

> We now take an entirely different view of the knowledge that the learner must acquire. Today's concern with both the semantic element in language and motivation requirements for successful learning goes a long way toward overriding some of those earlier simplistic warnings. (p. 4)
>
> For many educators selectivity was the answer to the made-for-ESL versus authentic song controversy. Pearse (1981) recognized that controlling language was often necessary for beginners, but argued that "with careful selection from the 'top twenty' and best selling LPs, this can be done quite easily" (p. 9). McBeath (1986) advised teachers to be selective when using songs with non-standard grammar or excessive slang, but maintained that songs especially constructed for ESL were often not as effective as authentic music since the lyrics could just end up becoming "a meaningless collection of phonemes" written to satisfy a narrow pedagogical objective (p. 44).[2]

Classroom Research into the Use of Music and Songs

Several studies measuring improvement in foreign language vocabulary and listening ability have shown that material based on songs is as effective as conventional material. Hahn (1972) found that junior high school males studying German vocabulary through songs achieved significantly higher scores on vocabulary tests than subjects studying vocabulary through dialogs. Medina (1993), in a study of Spanish-speaking elementary school subjects, found that exposure to English vocabulary through songs produced the same gains as through spoken versions of the material. A comparison of song-based listening texts with traditional dialog or narrative listening texts (Alley, 1990) found that first year high school Spanish learners made equally significant progress in improving listening comprehension with both types of material. Wilcox (1995) compared ESL learners' pronunciation development using songs and found that subjects progressed as much with songs as with nonmusical methods.

In a comparison of two different song methods, Grant, Clark and Koch (1996) found that students who studied listening through song-based gap-fill quizzes made the same progress as those who studied through song-based comprehension quizzes.

Student Attitudes Toward the Use of Songs in the Language Classroom

Kanel and Grant (1993) surveyed 550 Japanese college students who studied listening using popular song gap-fill quizzes. Regardless of English proficiency level, major, or sex, the respondents indicated that the song quizzes increased their interest and motivation in studying English, were as beneficial as, or more beneficial than, the other classroom materials in improving their listening ability, and felt that the exercises should be done often, perhaps every class. In a similar survey of American students studying Japanese (Jolly, 1975), the respondents indicated that songs created a relaxed and enjoyable atmosphere, livened up the pace of the lessons, and were an effective means of increasing vocabulary, studying Japanese culture, and discovering the relationship between language and culture.

To further examine the use of songs for language teaching, particularly listening development in an EFL teaching context, a study was conducted at a four-year university in Osaka, Japan, during the 1993-94 academic year. The researcher specifically set out to determine:

1) whether students' listening comprehension ability would progress as effectively with song-based tasks as it would with conventional nonmusical listening tasks, and
2) whether song-based tasks are equally effective at all levels of English proficiency.

The Study

Hypotheses

1) Scores obtained on a post-treatment test of listening comprehension for subjects using song-based tasks (song group) would not differ significantly from scores obtained by subjects using conventional nonmusical listening tasks (text group);
2) Scores obtained by subjects at each level of proficiency (A-D) in the song group would not differ significantly from scores obtained by subjects in the text group at the same level of proficiency;
3) Scores obtained for both song and text groups would show significant improvement over pre-test scores;
4) Scores obtained for subjects at all levels of listening proficiency in both groups would show significant improvement over pre-test scores.

Method

Subjects: The 692 subjects, native speakers of Japanese, were predominately male non-English majors enrolled in 20 first- or second-year required English classes at a Japanese university.

Design: Assignment to classes was based on students' academic majors, identification numbers, and year in school rather than on placement tests or self-grouping. The randomness of this assignment was deemed satisfactory for a quasi-experimental design. Ten full-time instructors, nine Japanese and one American, the researcher, who were teaching approximately seven classes each of this type were involved in the study. The instructors were directed to select two of their classes and assign one class to the song group and the other to the text group, resulting in 10 classes in each group. The instructors made an effort to select the two classes from the same academic major, and with the same course designations. (See Appendix A: Table of Instructors, Method, Majors, Year, and Texts).

The English curriculum consisted of four required courses: a) English I, first year reading; b) English II, first year conversation; c) English III, second year reading, and d) English IV, second year conversation. In addition to the class involved in the project, subjects were enrolled in one other English class during the academic year. First year subjects took both English I and II, and second year subjects took both English III and IV. Though no control over the amount of listening practice in their other classes was possible, it was reasoned that the randomization of class assignments to the study groups would balance the effect of any outside practice, in effect giving neither group an advantage.

Classroom materials and procedures: Students in the song groups were given listening practice through a series of song gap-fill worksheets prepared by the researcher (see Appendices B: List of Songs Used by the Song Groups, and C: Sample Song Worksheet). Gap-fill exercises were chosen over other types of listening tasks (e.g., dictation, true/false or multiple choice comprehension questions, passage correction, scrambled lyrics) because they are the simplest to construct and probably the most commonly used song-based tasks (Griffee, 1992; Murphey, 1992). More importantly, for low to intermediate EFL learners, those tasks provide practice in listening discrimination (i.e., distinguishing among L2 sounds, recognition of reduced forms, and word boundaries), which can lead to increased overall comprehension (see Richards, 1983; Rost, 1990).

Songs were chosen by the researcher based on experience using song gap-fill exercises during the two years prior to the current study (Kanel & Grant, 1993). Selection was based primarily on whether the songs had a) relatively clear enunciation and normal rhythm and intonation patterns, b) a conversational or narrative style, and c) a fairly wide range of vocabulary and grammatical structures. In all cases, taped versions of the songs by the original artists were used. Deletions were of a single lexical item except for occasional contracted forms. There were 15-30 deletions in each song depending on the song's tempo and length. The difficulty of the items was gradually increased over the course of the study as students became accustomed to the exercises. To further stimulate interest in the song exercises, a brief 120-150 word background of the singer(s) was provided for the instructor to read to the class before the quiz, if time permitted (see Appendix D: Sample Background Sheet).

Teachers played each song two times for the quiz, then put the answers on the blackboard while students corrected their own quizzes. Teachers then played the song for a third time while students analyzed their errors.

Students in the text groups were taught listening with nonmusical materials, either commercially available textbooks, or instructor created listening worksheets (see Appendix E: Sample of Instructor Prepared Cloze-Dictation Worksheet). The textbook selection was left to the discretion of individual instructors, however, the researcher provided a list of textbooks which contained exercises with items comparable those appearing on the JACET *Basic Listening Comprehension Test* (see discussion of *Measure* below). Instructors were directed to select materials which they felt would be most beneficial to students' listening comprehension development.

Listening materials in both text and song groups were used 20-30 minutes a week for the 14 class periods between the pre- and post-tests. For the remaining 60-70 minutes of class time, instructors taught equivalent lessons to both groups in accordance with the course designation. Instructors were directed to provide no additional listening practice during this time.

Measure: For both financial and practical considerations, the Japan Association of College English Teachers (JACET) Basic Listening Comprehension Test was selected as the most appropriate measure of non-English majors' English proficiency. Because there was no part B available at the time, it was used for both the pre- and post-test. Since the test administrations were separated by more than six months, and the answers never revealed to the subjects, it was reasoned that any practice effect would be minimal, and in any case equal for both groups. The pre-test was used to determine subjects' initial listening proficiency levels (i.e., A-D, with A being the highest level).[3] The post-test results were compared with the pre-test results to determine the effect of the two treatments.

The test consists of 40 multiple choice items divided equally into four sections: 1) picture, 2) statement, 3) dialog, and 4) narrative. The taped instructions are in Japanese, and the items are heard only once in English. The time of the test is approximately 45 minutes. Subjects were given the pre-test in the second week of May, with no prior warning, and given the post-test in the third week of November, again with no prior warning.[4]

Though the construct validity of the test has not yet been empirically demonstrated, considerable attention has been given to reliability: JACET determined a Cronbach Alpha value of approximately 0.9 (JACET, 1993).

Analyses

Analyses were based on the raw scores obtained on the JACET Test (maximum score = 40 points). Mean scores on the pre-test were compared with post-test scores using analysis of variance (ANOVA) to determine the effect of the two treatments. There was no previous support for positing a difference in mean scores between the two main treatment groups, thus, null hypotheses were adopted for hypotheses one and two. The significance level was set at $\alpha < .05$, non-directional.

For hypotheses three and four, it was felt that 14 class periods of listening practice with either treatment would result in enough improvement in listening comprehension ability to posit statistically significant gains in test scores. The significance level was set at $a < .05$, directional.

Results

Table 1 presents pre- and post-test descriptive statistics for the two treatments and proficiency levels.

Initially, histograms and cell plots of mean scores and standard deviations on the pre-test for the total population ($n = 692$), song group ($n = 334$), and text group ($n = 358$) were examined and showed normal distributions and variances, satisfying the assumptions for ANOVA. A two-way ANOVA (Table 2), using the pre-test scores as the dependent variable and treatment and level as the independent variables, showed no significant differences between the two main treatment groups ($F = .173$; $p = .6776$; $\alpha = .05$), and that both groups had four significantly different levels ($F = 1305.518$; $p = .0001$). This supports the earlier stipulation that assignment was random and that the two groups were at the same level of proficiency.

Table 1: Pre- and post-test descriptive statistics

Method	Level	*n*	Pre-test Mean	Pre-test *SD*	Pre-test Std. Error	Post-test Mean	Post-test *SD*	Post-test Std. Error	Gain
Song	all	334	17.308	4.568	.250	19.746	5.668	.310	+2.438
Text	all	358	17.693	5.323	.281	19.846	5.875	.310	+2.153
Song	A	14	27.571	1.555	.416	31.000	2.717	.726	+3.429
Text	A	30	28.133	1.697	.310	29.467	4.232	.773	+1.334
Song	B	96	21.740	2.084	.213	23.479	4.867	.497	+1.739
Text	B	100	22.020	1.954	.195	23.030	4.906	.491	+1.010
Song	C	177	15.853	1.868	.140	18.0181	4.086	.307	+2.328
Text	C	182	15.599	1.883	.140	17.747	4.306	.319	+2.148
Song	D	47	10.681	1.400	.204	14.660	3.766	.549	+3.979
Text	D	46	9.761	1.980	.292	14.957	2.875	.424	+5.196

A three-way repeated measures ANOVA (Table 3), using pre- and post-test scores as the dependent variable and treatment, level and pre- and post-test scores (the repeated measures) as independent variables showed no effect for treatment on the two main groups ($F = 1.089$; $p = .2971$). This confirmed the first null hypothesis of no significant difference between treatments. Interaction tests on the effect of level and treatment confirmed the second null hypothesis of no effect for treatment on proficiency levels ($F = 1.920$; $p = .1249$). The third hypothesis, that scores for both treatment groups would show significant improvement from pre- to post-test, was confirmed ($F = 152.641$; $p = .0001$), as was the fourth hypothesis that sub-

Table 2: Two-way ANOVA of the pre-test

Source	*df*	Sum of Squares	Mean Square	*F* Value	*p* Value
Treatment	1	.614	.614	.173	.6776
Level	3	13901.630	4633.877	1305.518	.0001
Treatment * Level	3	28.666	9.555	2.692	.0453
Residual	684	2427.828	3.549		

Dependent: Pre-test

Table 3: Three-way repeated-measures ANCOVA of the pre- & post-tests

Source	*df*	Sum of Squares	Mean Square	*F* Value	*p* Value
Treatment	1	16.748	16.748	1.219	.2700
Level	3	123592.008	7864.003	572.212	.0001
Treament * Level	3	5.320	1.773	.129	.9429
Subject (Group)	684	9400.327	13.743		
Pre- Post Test	1	1248.549	1248.549	152.641	.0001
Pre- Post Test * Treatment	1	8.905	8.905	1.089	.2971
Pre- Post Test * Level	3	327.783	109.261	13.358	.0001
Pre- Post Test * Treatment * Level	3	47.119	15.706	1.920	.1249
Pre- Post Test * Subject (Group)	684	5594.891	8.180		

Dependent: Measure

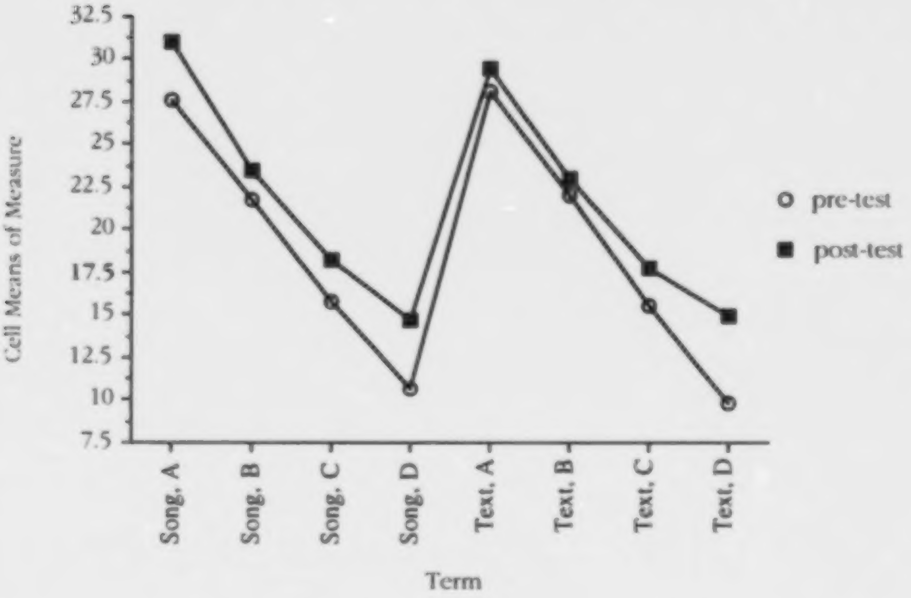
jects at all proficiency levels in both treatments groups would improve significantly ($F = 13.358$; $p = .0001$).

An interaction plot illustrating the progress of both treatments at all four levels from pre- to post-test is shown in Figure 1.

Follow-up Questionnaire

The responses of the two groups to a follow-up questionnaire varied little in regard to evaluation of the procedures used in class, or the positive educational benefits of their respective listening exercises. Questionnaire items which asked whether the exercises helped improve pronunciation, intonation, and contracted forms all received approval ratings of 50-60% by both groups. In addition, both groups indicated that they wanted to do the exercises regularly. The song group, however, favored doing the exercises every week significantly more than the text group (song = 74%; text = 58%; Chi Square = 15.504; $p = .0001$) and found more value in the time spent on the quizzes (song = 61%; text = 48%; Chi Square = 9.686; $p = .0018$). The song group also expressed significantly more interest in studying English than the text group (song = 50%; text = 32%; Chi Square = 20.008; $p = .0001$).

Figure 1: Interaction Plot Effect:
Pre- & Post-Test Treatment, Level Dependent Measures



Discussion

The specific concerns of this study were to determine whether song-based tasks, in the form of regular gap-fill quizzes, would be as effective at improving listening comprehension as conventional listening tasks at all levels of listening proficiency as operationally defined by the JACET Basic Listening Comprehension Test. The post-test scores showed significant improvement for both text and song groups at all levels and that neither treatment was more effective than the other. These findings confirmed all of the original hypotheses and concurred with the studies discussed earlier that suggest song-based language teaching tasks are as effective as nonmusical tasks.

For the purposes of this study, only song gap-fill tasks were utilized for the song treatment. It was reasoned that the relatively low level of subjects in this study would benefit most by practice in listening discrimination, and that improvement in the ability to distinguish words in context would result in increased listening comprehension ability. The results appear to support this reasoning. The materials used by the text group included listening discrimination tasks such as gap-fill and dictation, as well as comprehension items generally more similar to those found on the JACET Listening Test than the song group's gap-fill tasks. Although the song group performed as well as the text group on the final measure of listening comprehension, a combination of gap-fill tasks and comprehension questions based on the songs' contents may have increased the song group's scores enough to have made them statistically significant. In addition, it is possible that the motivational advantages posited for song-based tasks could not compensate for the difficulty lower proficiency subjects had with the speed, vocabulary, and abstract and poetic nature of songs. In fact, though the number of higher proficiency subjects in this study was small, one might conclude by looking at their mean scores and the interaction plot that higher proficiency students do better with songs.

Conclusion

The number of intervening variables present in a study of this nature, (i.e., outside exposure to English, methods of individual instructors, gender, time of day, classroom environment, seating arrangement, musical training, and aptitude) make it impossible to posit a direct causal relationship between the two methods and the improvement in listening ability. Tighter controls on these variables in future studies could reveal what specific advantages songs and music might have over non-musical listening tasks. Although the results should be considered more descriptive than inferential, teachers may interpret these findings as support for rejecting the idea that song use in the L2 classroom is limited to entertainment or mood enhancement and has little practical value. Moreover, it is apparent from students' responses to the follow-up questionnaire that they feel song listening exercises are beneficial and want to do them regularly.

It is likely that the use of music and songs in L2 classrooms would gain wider acceptance if there were more empirical research demonstrating positive effects on second language acquisition. Although this study concerned itself with listening ability, improvement in students' oral production (i.e., intonation, pronunciation, stress, vocabulary, and contracted forms) through study with music and songs needs further investigation. Studies measuring the gains achieved through use of song-based activities compared with gains made through conventional activities are needed in areas such as the retention of specific forms, and the application and accuracy of their use. In addition, further research into the cognitive and affective advantages of music and songs is necessary, particularly in the areas of the Din and SSIMH phenomena, right-brain/

left-brain theory, and the relationship between musical ability and language aptitude.

Kim Kanel is an associate professor in the Department of General Education at Kinki University, Osaka. He has taught EFL in Japan for 18 years. His research interests include the use of media and peer-interaction in language teaching and sociolinguistics.

Notes

1. See Murphey (1990a) *Song and Music in Language Learning*, and *Music and Song* (1992b) for lists of both ELT song books and books with activities to teach English through authentic songs.
2. For further discussion, see Kanel (1996, pp. 118-120).
3. Originally there were five levels assigned by the JACET test, S, A, B, C, D, with S being the highest. However, because there were only 14 subjects in the S category in the subject population, and their scores were so much higher than the other subjects in their treatment groups, their data was deleted from the database.
4. Since the JACET test is available only twice a year, May and November, the amount of time for treatment was limited to 14 weeks, not including the intervening two month summer break.

References

Alley, D. (1990). *The role of music in the teaching of listening comprehension in Spanish*. University Microfilms International Order No. 8823758. (RSN 1113510).

Barber, E.J.W. (1980). Language acquisition and applied linguistics. *ADFL Bulletin, 12*, 26-32.

Berglund, J. (1983). *The Beatles once and forever*. Kyoto: Chuotosho.

Botari, S., & Evans, J. (1982). Effects of musical context, type of vocal presentation, and time on the verbal retention abilities of visual-spatially oriented and verbally oriented learning disabled children. *Journal of School Psychology, 20*(4), 324-338.

Coe, N. (1972). What use are songs in FL teaching? *International Review of Applied Linguistics in Language Teaching, 10*(4), 357-360.

Cranmer, D. & Laroy, C. (1992). *Musical openings*. Essex, UK: Longman.

DeSelms, C. (1983). Music to learn by. *JALT Newsletter, 7*(4), 6-7.

Dougill, J. (1989). *Rock classics: A study of rock singers and lyrics*. Tokyo: Macmillan Language House.

Dubin, F. (1974). Pop, rock and folk music: An overlooked resource. *English Teaching Forum, 12*(3), 1-5.

Gfeller, K. (1983). Musical mnemonics as an aid to retention with normal and learning disabled students. *Journal of Music Therapy, 20*(4), 179-189.

Grant, P.B., Clark, C., & Koch, T. (1996). Listening comprehension through songs in the English language classroom: A comparison of cloze and question exercises. *The Research Journal of the Department of Teacher Education, Kinki University, 7*(2), 77-97.

Griffee, D.T. (1992). *Songs in action*. London: Prentice Hall International.

Hahn, S.M. (1972). *The effect of music in the learning and retention of lexical items in German*. Education Resources Information Center (ERIC Document # ED 119455).

House, J., & Manning, J. (1992). *Say it in song*. Tokyo: Macmillan Language House.

Isern, B. (1958). The influence of music upon the memory of mentally retarded children. *Music Therapy, 18*, 162-165.

JACET (1993). *Manual for JACET basic listening comprehension test*. Tokyo: Kaitakusha.

Jalongo, M., & Bromley, K. (1984). Developing linguistic competence through song picture books. *Reading Teacher, 37*(9), 840-845.

Jolly, Y. S. (1975). The use of songs in teaching foreign languages. *Modern Language Journal, 59*(1), 11-14.

Kanel, K. & Grant, P. (1993). The use of cloze song quizzes in the college classroom: A survey of student attitudes. *The Research Journal of the Department of Liberal Arts, Kinki University, 24*(3), 25-72.

Kanel, K. (1995). *Pop song listening*. Tokyo: Seibido.

Kanel, K. (1997). *Enjoy pop songs*. Tokyo: Seibido.

Kanzaki, H. (1988). *A shorter course in pops*. Tokyo: Nan'un-do.

Krashen, S.D. (1983). The din in the head, input, and the language acquisition device. *Foreign Language Annals, 16*, 41-44.

Loew, H.Z. (1979). Tuning in: Popular music in the second language classroom. *Foreign Language Annals, 12*, 271-274.

Lozanov, G. (1979). *Suggestology and outlines of suggestopedy*. New York: Gordon and Breach Science Publishers.

McCarthy, W. (1985). Promoting language development through music. *Academic Therapy, 21*(2), 237-242.

Medina, S.L. (1993). The effect of music on second language vocabulary acquisition. FLES News: National Network for Early Language Learning, 6(3), 6-8 (ERIC Document # ED 352-834).

Meyer, L.B. (1956). *Emotion and meaning in music*. Chicago: The University of Chicago Press.

Mosdell, C. (1984). *The Beatles: A modern way to learn English through the songs of the Beatles*. Tokyo: Kinseido.

Murphey, T. (1984). English language music: The ESL of Swiss adolescents. *English Teachers Association Switzerland Newsletter, 3*, 18-19.

Murphey, T. (1988). The discourse of pop song lyrics. *IATEFL Newsletter, 100*(8), 36-37.

Murphey, T. (1989). The where, when and who of pop lyrics: The listener's prerogative. *Popular Music, 8*(2), 185-193.

Murphey, T. (1990a). The song stuck in my head phenomenon: A melodic din in the LAD? *System, 18*(1), 53-64.

Murphey, T. (1990b). *Song and music in language learning: An analysis of pop song lyrics and the use of music and song in teaching English to speakers of other languages*. Bern, Switzerland: Peter Lang.

Murphey, T. (1992). *Music and song*. Oxford: Oxford University Press.

Parr, P.C. & Krashen, S.D. (1986). Involuntary rehearsal of second language in beginning and advanced performers. *System, 14*, 275-278.

Posener, A. (1987). *The Beatles story*. Tokyo: Macmillan Language House.

Richards, J. (1969). Songs in language learning. *TESOL Quarterly, 3*,161-174.

Richards, J. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly, 17*(2), 219-240.

Rosenfeld, A. H. (1985). Music the beautiful disturber. *Psychology Today*, December, 48-56.

Rost, M. (1990). *Listening in language learning*. London: Longman.

Sato, T., & Sasanuma, M. (1988). *Let's take a singing break*. Tokyo: New Currents International.

Schuster, D., & Mouzon, D. (1982). Music and vocabulary learning. *Journal of the Society for Accelerative Learning and Teaching, 7*(1), 82-106.

Sekara, A.T. (1985). Sound sense. *English Teaching Forum, 23*, 26-30.

Smith, L.E. (1976). The use of popular songs in an ESOL class. *English Language Teaching Forum, 14*(1),1-6.

Someya, M., & Ferrasci, F. (1988). *Great American hit songs*. Tokyo: Macmillan Language House.

Stoudenmire, J. (1975). A comparison of muscle relaxation training and music in the reduction of state and trait anxiety. *Journal of Clinical Psychology, 33*, 490-492.

Urbancic, A., & Vixmuller, J. (1981). Using popular music in the foreign language classroom. *Canadian Modern Language Review, 38*(1), 81-87.

Wilcox, W.B. (1995). *Music cues from classroom singing for language acquisition: Prosodic memory for pronunciation of target vocabulary by adult non-native English speakers*. University of Kansas dissertation, May 1995. Dissertation Abstracts On-line.

(Received December 21, 1996; revised April 24, 1997)

Appendix A: Instructors, Song/Text Class Majors, Number of Subjects, Year in School, and Text Class Listening Materials

Instructor	Song class major	*N*	Year	Text class major	*N*	Year	Text class listening material
Ashihara	Management	34	1	Management	36	1	*10 Minute Hearing*
Fujiwara	Business	39	2	Business	38	2	*10 Minute Hearing*
Kanazawa	Business Law	23	2	Management	40	2	*10 Minute Listening*
Researcher	Business	41	2	Engineering	41	2	*10 Minute Listening*
Kawanishi	Management	15	1	Business	28	1	*10 Minute Listening*
Kimura	Engineering	40	2	Economics	29	2	Exercises from video tapescript
Kusumoto	Engineering	29	1	Engineering	33	1	Listening lessons from course text
Okuda	Engineering	30	2	Engineering	42	2	Instructor created worksheets
Yamamoto, E.	Management	45	1	Management	36	1	Instructor created worksheets
Yamamoto, T.	Engineering	38	1	Engineering	35	1	Instructor created worksheets
Total		334		Total	358		

Appendix B: List of Songs Used by the Song Groups

	Title	Artist
1.	*Love Me Tender*	Elvis Presley
2.	*Imagine*	John Lennon
3.	*Stand By Me*	Ben E. King
4.	*Take Me Home Country Roads*	John Denver
5.	*Yesterday Once More*	The Carpenters
6.	*Wonderful Tonight*	Eric Clapton
7.	*Oh, Pretty Woman*	Roy Orbison
8.	*Tom's Diner*	Suzanne Vega
9.	*Honesty*	Billy Joel
10.	*Tears In Heaven*	Eric Clapton
11.	*I Will Always Love You*	Dolly Parton
12.	*Help!*	The Beatles
13.	*Time In A Bottle*	Jim Croce
14.	*Are You Lonesome Tonight?*	Elvis Presley

Appendix C: Sample Song Worksheet

Love Me Tender by Elvis Presley

Love me tender, love me (1)____________.
Never let me (2)______________.
You have made my life (3)______________,
And I love you (4)______________.
* Love me tender, love me (5)______________.
All my (6)______________ fulfill.
For my (7)______________ I love you.
And I (8)______________ will.
Love me tender, love me (9)______________.
(10)______________ me to your heart.
For it's (11)______________ that I (12)______________,
And we'll never (13) ______________
Repeat *
Love me tender, love me (14)______________,
Tell me you are (15)______________.
I'll be yours through all the (16)______________,
Till the end of (17)______________.
Repeat *

Appendix D: Sample Background Sheet

Background: Elvis Presley

Elvis Presley, born in Mississippi in 1935, got his first guitar for his eleventh birthday. In his teens he listened to white country & western music and black rhythm & blues. His early singing style, called 'rock-a-billy,' combined both these types of music, and many listeners thought he was black at first. He worked as a truck driver until he signed a recording contract in 1954, and sang his first big hit, *Heartbreak Hotel*, on the *Ed Sullivan Show* in 1956. After two years in the army, from 1958 to 1960, Elvis appeared in movies during the 60s and 70s. He died of a heart attack in 1977 at the age of 42. *Love Me Tender*, from the movie of the same title, was recorded by Elvis in 1956.

Appendix E: Instructor Prepared Cloze-Dictation Worksheet

Listening Quiz (2)
Student no. ( ) Name ( )

Biological Clocks

Every (1) thing has what scientists call a biological clock that controls behavior. The biological clock tells plants when to form flowers and when the flowers should (2). It tells insects when to (3) the protective cocoon and fly away. And it tells animals and human beings when to (4), sleep and wake. It controls our (5) temperatures, the release of some hormones and even dreams. (6) outside the plant and animal affect the actions of some biological clocks. Scientists recently found, for example, that a (7) animal called the Siberian hamster changes the color of its fur because of the (8) of hours of daylight. In the short days of winter, its fur becomes (9). The fur becomes gray-brown in (10) in the longer hours of daylight in summer. Inner signals (11) other biological clocks. West German scientists found that some (12) of internal clock seems to order birds to (13) their long migration flights two times each year. Birds prevented from flying become restless when it is time for the (14). But they become (15) again when the time of the flight has ended.

Learning Pronunciation and Intonation of Japanese through Drama by Beginning Language Students: A Case for Reflective Journals

Harumi Moore
Australian National University

This paper portrays the benefits of using reflective journals in a tertiary education environment. It focuses its discussion on how the use of a reflective journal brought learners towards a closer approximation of native-like pronunciation and intonation, which was one of the objectives of a drama component in a first year Japanese program at the Australian National University in 1995. The use of the journal enhanced learner consciousness in cognitive and metacognitive learning, and serving as an excellent resource for qualitative research, and enhanced teachers' readiness and their ability to identify and analyse many learning issues. Above all, it fostered empathy among teachers towards students' learning experiences, and developed a sense of a cooperative relationship between students and teachers as co-participants in the learning process.

この論文は、高等教育において日記による内省を使うことの利点を記述する。1995年、オーストラリア国立大学の日本語プログラムの一年めの授業の一部であったドラマの目的の一つは、母語話者のような発音とイントネーションにより近づくことであった。本稿は、そのための学習過程が日記による内省によって、どのように促進されたかを論じる。日記による内省は、学習の認知的およびメタ認知的側面に関する意識を高め、質的研究のすばらしいリソースとなり、また、教師が学習の問題を特定し、分析するレディネスと能力を高めた。さらに、教師の間に学生の学習経験に対する共感をうみだし、学習のプロセスに共に参加する者として、学生と教師の間に協力的関係を育てた。

Students who study Japanese at the Japan Centre, Australian National University, have mixed backgrounds in relation to previous exposure to Japanese.[1] Expectations held by both students and teachers were that oral fluency, particularly native-like fluency of pronunciation and intonation, could not be acquired just through language classes at a university outside Japan, where contact hours are lmited, and opportunity for exposure to the language once one walks out of the classroom is small.

JALT Journal, Vol. 19, No. 2, November, 1997

As part of a major curriculum development project, a drama component was introduced into a first year Japanese course in 1995. The rationale of the drama component was multifold: exploring various educational objectives of foreign language learning such as helping students learn a language in cooperative relationships with peers and with the teacher; exposing students to authentic spoken language models from an early stage; challenging students to learn the language without being analytical about every detail of the structure of the language such as the morphological or syntactic structures of a phrase, and teaching body language. In particular, the drama component enabled us to investigate how accurately students could learn the pronunciation and intonation of Japanese through immersing themselves in an intensive and repetitive process of listening to and repeating pieces of language and rote memorization.

Part of this component was the use of student-kept reflective journals. Teachers hoped that keeping reflective journals would promote students' critical thinking skills and enhance their awareness of the learning of pronunciation and intonation. It was hoped that the positive effects of journal-keeping would compensate for the external disadvantage of restricted time for formal classroom interaction and the lack of regular contact with Japanese speakers outside the classroom.

Before describing the progress which was monitored and enhanced through the use of reflective journals, I will first discuss three key issues which form the background to the present study, i.e. the use of reflective journals as a tool for enhancing learner awareness, the mastery of native-like pronunciation and intonation by adults, and the benefits of drama as part of a second language teaching curriculum.

Learner Awareness and Journal Keeping

In education in general, developing autonomous learning or taking control of one's learning has been advocated for some time:

> [M]any practitioners throughout the world are trying to establish ways in which they can assist students to become less dependent upon them as teachers and to design courses which involve students more deeply in learning and in making decisions about what they will study. (Boud, 1986, p. 21)

To promote autonomy, teachers and educators have placed increasing stress on observing the process of learning from the learner's point of view, so that teachers can help learners enhance their awareness and take control of their own learning. Foreign language learning is not an exception:

> There is by now a substantial body of research outlining the behaviours learners use and describing the thought processes they engender while learning a foreign or second language. In particular, the focus of research has been on identifying the behaviours and thought processes used by language students to learn a foreign or second language. (Rubin 1987, p. 15)

The growing interest in the study of the benefits of diary-keeping is reflected in the increasing body of literature which not only addresses educational and interpersonal development benefits but also the benefits for teacher education and for understanding the social and cultural norms of students and teachers (Bailey, 1983; Bailey, 1990; Kreeft-Peyton & Reed, 1990; Matsumoto, 1987; Peyton, 1990; Schumann, 1980; Staton, 1987; Staton, Shuy, Peyton & Reed, 1988).

Keeping a learner diary is beneficial for raising awareness of the learning process (Bailey, 1990, p. 223-224) because to a certain extent it helps students to monitor and assess themselves. A learner diary can be used simply as a record-keeping tool. Carver and Dickinson suggest that a learner diary should contain entries such as "Date; Lesson in text book; Main activities; How I performed; What difficulties I had; What difficulties I still have; and What I intend to do next" (cited in Dickinson, 1987, p. 185). Part of Oskarsson's proposed detailed form of a learner diary includes a self-assessment section as well (cited in Dickinson, 1987, p. 186). Learner diaries can be used to explore learning strategies: "the writing of a diary helped her 'evaluate her own learning strategies, enabling her in some cases to manipulate strategies so that she received the most benefit'" (Henze, cited in Rubin, 1987, p. 16). Dickinson (1987) also states that a learner diary is beneficial when used during consultations with the teacher, particularly when the learner is not fully autonomous (p. 185).

Further, the reflective journal, when used to promote dialogue between the student and the teacher, can be a powerful tool from the point of view of qualitative action research, enabling teachers to develop empathy with students. As Bogdan and Biklan (1982) state, "the goal is to understand the subjects' world and to determine how and with what criteria they judge it" (p. 210). Diaries are beneficial as a second language classroom research tool. (see Allwright, 1983; Allwright & Bailey, 1991; Chaudron, 1988; Richards & Lockhard, 1994)

Because reflective journals act as a direct communication channel between individual students and teachers, they help to deepen the student-teacher relationship and develop empathy. The original purpose of introducing diary keeping in education was for "better personal communication and mutual understanding between each individual student and teacher" (Staton, 1987, p. 157).

The initial primary purpose for incorporating the reflective journal in the Japan Centre course was educational, to heighten learner awareness. If we are to aim for the maximum desired language proficiency in a foreign language learning situation, especially in a country where the target language is not spoken and contact hours are limited, it is even more important to promote conscious and autonomous learning habits. Rubin (1987, p. 17) states "it is essential for students to be able to take control of their own learning process so that they can learn outside the classroom once they are on their own." It is for this reason that reflective journal keeping was incorporated into a Japanese language course at the Australian National University.

Pronunciation

Imperfect mastery of pronunciation and intonation of a second language is heard as a "foreign accent." There appear to be multiple factors contributing to the pronunciation and intonation attained by the second language learner. Some claim that the age at which one starts learning the second language is a crucial factor (Asher & Garcia, 1969; Krashen, Long & Scarcella 1979; Oyama, 1976; Patkowski, 1980; Scovel, 1988; Seliger, 1978; Seliger, Krashen & Ladeford, 1975), which supports a theory of the existence of a critical period (Lennenberg, 1967) or a sensitive period (Lamendella, 1977). Length of residence in a place where the target language is spoken is also suggested to be a factor (Purcell & Suter, 1980).

However, some studies claim that older people are not disadvantaged in the attainment of native-like pronunciation of the second language (e.g. Jones, 1985; Neufeld, 1978; Olson & Samuels, 1973; Snow & Hoefnagel-Höle, 1977). Furthermore, some writers suggest that there are personal factors that are relevant to the degree of attainment of native-like pronunciation and intonation including: difference in individual aptitudes such as "phonetic coding ability" (Carroll, 1981, p. 105); a capacity to mimic sounds in a foreign language (Purcell & Suter, 1980); motivation to pronounce the target language accurately (Purcell & Suter, 1980); the degree of "empathetic capacity," i.e. the more empathic the learner is towards the target language speech community the more likely the attainment of native-like pronunciation (Guiora, Beit-

Hallahmi, Brannon, Dull & Scovel, 1972); the cultural background of the learner (see Paulston, 1978; Busch, 1982), and the degree of phonological interference of the native language (see Odlin, 1989; Purcell & Suter, 1980; Suter, 1976).

Drama

Drama has not only been incorporated in education in general, it has also proved its benefits in second language teaching. Stern (1980, pp. 78-82) argues that drama encourages psychological factors in the learner which help develop communicative competence in the second language, such as enhancing "motivation," "self-esteem," reducing or eliminating "sensitivity to rejection," and increasing capacity for "empathy." Some pedagogical materials have been written for using drama techniques for second language learning (see Holden, 1981; Maley & Duff, 1982; Wessels, 1987). While factors relevant to the learning of pronunciation and intonation of the second language are being studied, there is also a pedagogical interest in the *methods* of teaching pronunciation and intonation (see Brown, 1992; Morley, 1994; Tench, 1981; Wong, 1987). Using drama for teaching pronunciation is one of those methods. Stern (1980) states that dramatics were effective in speech therapy in children for psychological reasons. In a chapter on techniques for improving pronunciation using drama, Wessels (1987, pp. 62) states "speech is more than simply repeating what you hear . . . the shape of the mouth, posture, the mechanics of breathing, and even facial expressions are part and parcel of correct pronunciation."

With these key issues as background, the teachers involved in the present study hoped that rote memorization required by the drama class would help first year students who had not lived in Japan to achieve optimal approximation of native-like pronunciation and intonation, at the same time using reflective journals to enhance the learning process.

The Study

The present study was conducted in the drama component of Spoken Japanese 1, a five-hour-per-week, semester-long unit of the first year Japanese program at the Japan Centre, Australian National University, in 1995. The drama component was one-and-a-half hours in length per week, and was introduced along with the other components as part of an experimental curriculum development project.

Method

Subjects: Subjects were 52 first year students enrolled in the Japanese course Spoken 1. The mother tongue of most was English. There were eight non-native speakers of English, from Korea and Hong Kong. The group included 32 females and 20 males. The subjects were divided into six groups of 2 to 14 students. (The uneven group sizes resulted from other student commitments.) Most subjects had some experience learning elementary Japanese during junior and senior high school. However, a few had no previous knowledge of Japanese.

Teachers: Three teachers were involved in the designing and planning of the drama component. Four teachers, all native speakers of Japanese, did the actual teaching.

Materials: Four different scenarios, short plays, were written by teaching staff at the Japan Centre, aiming for authenticity of colloquial expressions, male and female speech patterns, interruptions, and unfinished sentences. Each play was 12 to 15 minutes in length when acted out. The scenarios differed from dialogues contained in beginners' textbooks, which are usually heavily graded, employ restricted vocabulary and grammar, and use predominantly short sentences. The scenarios were also written so that the difference in the amount of lines to be acted out by each character would help bridge the gap between those who had studied Japanese before and complete beginners. Each script required a different number of roles, so that one could be chosen to match the size of the class. The classroom teachers chose a script in the first meeting according to the number of students enrolled. Students were given scripts in Japanese, along with translations into natural English. The Japanese scripts were written with *kana* and *kanji*. This was intended to give students a visual guide to the word boundaries associated with intonation patterns. Students were not expected to be able to read the *kanji*.

Audio tapes modeled the scripts at a natural speed. Students purchased these for out-of-class preparation.

Procedure: The first half of the semester was used to study the script, paying attention to detail. Students studied the script section by section in class, listening to the teacher and the tape. The teacher gave feedback and correction to the previous week's out-of-class preparation by individual students in class. Time was spent to discuss issues brought up in the journals. In the early stages, reading lines aloud in English was used to help students get into the character and the mood of their lines as well as to enhance group dynamics. The latter half of the semester was used to put the play together in Japanese. Students directed the play themselves.

Assessment: Two formal assessments were conducted. One was done in Week 8. Teachers felt that it was necessary to give a formal assessment halfway through the semester to encourage students to memorise their lines as well as to learn them with accurate pronunciation and intonation. In the Week 8 assessment, each group was asked to recite part of the script. Although the recitation was performed by the group, students were individually assessed for the degree of accuracy in pronunciation, intonation, and line memorization. The assessment recitation was recorded on audio tapes. Two teachers attended the assessment session for each group. Students were given cue words by one of the teachers when they could not remember a line.

The second assessment was the final performance in Week 13. Pronunciation, intonation, and memorisation of lines were assessed, and additional points given for good acting. Students used props and prepared simple costumes. Teachers helped make some props. All first year students were invited to watch other groups perform. Many also invited friends and families. Teachers at the Japan Centre not directly involved in the drama classes were also invited. The performances were videotaped.

The first assessment constituted 40% of the drama component, and the final assessment 60%. Difficulties experienced with this assessment method are discussed below.

Conscientious completion of journals was encouraged by grading journal-keeping as part of the assessment of the overall five-hour-per-week course.

Reflective journal use: Students were asked to write and submit a weekly (later bi-weekly) journal in English. Journal sheets asked open-ended questions which often addressed issues not only relating directly to pronunciation and intonation but also to learning strategies (see O'Malley & Chamot, 1990; Oxford, 1990; Rubin, 1975, 1981; Rubin & Wenden, 1987; Wenden, 1985). In order to enhance metacognitive strategies, i.e. strategies used to "oversee, regulate or self-direct language learning" (Rubin, 1987, p. 25), teachers encouraged students to "plan," "monitor," and "evaluate" through their journals the overall learning process, including time management, and resources and strategies used. Use of cognitive strategies was also consciously explored and identified in journal questions. Such strategies included "repetition," "rehearsal," "experimentation," "imitation," "attention to detail," "memorization," and "direct analysis" (Rubin, 1987, p. 25). The weekly journal questions were designed as an action research tool, in which issues were addressed as they became relevant. (See Appendix for sample journal questions.)

The Reflective Journals

The reflective journals maintained by the students provide a great deal of information about the learning processes of students studying Japanese as a foreign language. These journals were analysed by the researcher.

The first class was carefully prepared in order to produce a relaxed induction session, so that students could embark on the project with positive, confident attitudes and a clear objective in mind. This session focused on helping students become aware that they could apply their existing knowledge, skills, and experiences to achieve the task set by the drama project, and could take the initiative in this and other learning experiences. The in-class activities included forming a performing group, group reading of the English script, talking about the play, and visualising the characters and scenes. Some students also talked about their individual experiences and the strategies they had previously used, including acting, mimicking other people's speech characteristics, and memorising substantial lengths of text. The feedback from students in the pre-project journal was that they experienced a mixture of excitement and fear. Some comments from students' journals were:

> "Excited, but a bit scared."
>
> "It's gonna be enjoyable."
>
> "It's good to study Japanese in different ways."
>
> "Surprised, but it could be fun."
>
> "Surprised, and not looking forward to the drama."

Reading of the English script was repeated during the first few weeks. This helped students become familiar with the development of the story and with their role from an early stage. It was intended that this strategy would help students learn their lines in Japanese as meaningful phrases, rather than having them concentrate unduly on the syntactic construction of the authentic and therefore uncontrolled text. As students were able to read fluently in English, this process facilitated the process of forming group spirit:

> "Acting in English put us on the same level, without having to stumble over lines."
>
> "Acting the play in English helps a lot because we get to know the play as well as the other members. This makes it easier when it comes to the Japanese reading to know what sort of expressions to put into our voice."

Without this process, a certain amount of uneasiness would have been expected from students with no background in Japanese when working with students who had studied Japanese before.

Becoming Aware of "Intonation"

From an early stage, students commented in journals about the learning process, especially noting how they came to know something new by paying detailed attention to it. Discovering "intonation" was one of these things. In general, Japanese intonation is not taught in secondary school language classes, mainly because of lack of training, and thus lack of awareness, on the part of teachers. Those students who had studied Japanese before became aware of the rhythm of Japanese intonation for the first time, and the different stress and timing patterns than English. Many students had been imposing English intonation patterns on Japanese speech until then. For example, in their journals, two students said:

> "It was difficult to imitate the native's pronunciation and intonation. I had never thought about this aspect of speaking in Japanese before."
>
> "Intonation was almost as hard as pronunciation."

One student expressed that he felt embarrassed to try out the Japanese intonation pattern:

> "It sounded silly using different emphasis."

This comment lead to a class discussion of what it means to speak a foreign language. Students were assured that there was no need to feel embarrassed, and that the Drama class was an environment in which students were *pretending to be Japanese.*

Uncertainty about whether one can acquire native-like pronunciation and intonation in a foreign language was felt by some students at the outset. They seemed to assume that the ability might be something innate:

> "I have very little sense of my own voice. I'll just have to practice listening to it." (This student was a complete beginner but acquired quite an acceptable pronunciation and intonation.)

> "Intonation is okay if you have a 'musical ear'."

The fact that most students were able to imitate the native Japanese teacher in class, however, provided evidence that it may be possible for most students to develop this aspect of the language to a certain extent.

Exploring Strategies

Throughout the process, students explored different methods of achieving better pronunciation and intonation. The earliest sign of their realization that different strategies may yield different outcomes was experienced in one of the classes in Week 3. The class started off with students mimicking the teacher's pronunciation and intonation line by line. Everyone had a corresponding English script. Students were copying the teacher very accurately. The teacher was aware of the native-like pronunciation and intonation being produced. At this stage, students had not been given the Japanese script. The class was asked how they felt about this blind mimicking, and most of them were comfortable with it. However, some students expressed uneasiness at not being able to see what they were saying in Japanese. At this point, students were given the Japanese script. Practice continued, and students who had some knowledge of Japanese started to look at the Japanese script while repeating the lines after the teacher. Immediately, the teacher noted a drastic drop in the accuracy of intonation. The following are some of the entries from that week's journal:

> "I think that repeating the sentences after the teacher for pronunciation and production is fantastic. I felt I got more out of it when we didn't have the script in front of us—because as soon as the script was in front of me I began to have problems . . . and couldn't take in the intonation as well as before."
>
> "I think it was better going through it without having the Japanese script to distract us. It was more beneficial just blindly mimicking the intonation."

It was generally felt that from an early stage "listening to the model" was the most effective method to help the acquisition of accurate pronunciation and intonation, and most students proposed that they would use the model tape as much as possible for out-of-class preparation, planning on:

> " . . . listening to the tape to get the pronunciation and intonation."
>
> "Listening to the tape and subconsciously becoming aware and used to the natural tone/pattern of the language."

Many students realized that "repetition" was very important from an early stage:

> "I listened over and over again to the tape. I think it's working . . . . Repetition is my biggest help."

> "I listen to the tape and repeat with it at the same time. I find that this is perhaps the best method, as I can repeat the rise and fall of intonation. Constant repetition also helps."

One student drew intonation lines in the script as visual cues for the rise and fall of intonation:

> "I drew _ and – symbols for the intonation. This helped. . . ."

The fact that there were many lines to memorize demanded that students use the script for this task. But how they incorporated the audio and visual resources differed from one student to another:

> "Listening to the tape as well as reading the script was usually a very effective method for me."

> "I go through the script at least once, almost every day. I read my parts out aloud which I find works. I also listen to the tape while looking at the script."

Sometimes students used the script and the tape simultaneously, listening to the tape while visually following the lines in the script. Sometimes students used only one of the resources, reading aloud, for example, while looking at the script.

Differences in Learning Styles

Students used to analytical modes of learning experienced uneasiness in learning their lines solely by mimicking without understanding the detailed grammatical constructions. This problem was anticipated, because, in the interest of authenticity, the script was not graded for beginners. The journal question asked, "How do you feel about learning lines which are beyond your grammatical knowledge?" Some students expressed frustration:

> "I'm finding it quite hard, because it's hard to memorize something you can't really understand or know the full meaning."

> "I don't like saying things I don't understand."

On the other hand, some students were happy with the task:

> "It doesn't really bother me. Just to know them and what they mean will be great satisfaction."

> "It's not too hard, because I have a general understanding of what I'm saying. So it doesn't bother me."

The following entry illustrates the determination students needed to succeed in achieving the objective of the drama class of learning the lines and acting them with native-like pronunciation and intonation. This student had an analytical approach to language, but consciously challenged herself to use a method which she felt she needed:

> "I decided not to get stressed out if I didn't understand the grammatical structure–but it doesn't really bother me now as I can still learn my lines without this knowledge. I probably will go over the script and look at the grammar though, now that I'm more familiar with it." [This student had no prior knowledge of Japanese but with persistent effort achieved excellent pronunciation and intonation.]

Each of the six classes discussed the differences in learning styles among learners. They confirmed what Rubin (1987) points out:

> [G]iven the same learning environment, the same target language, the same native language, and the same language level, some learners will be more analytic in their approach to the learning task while others will be more intuitive; some learners will prefer to use written materials to access a foreign language while other will prefer to hear the language. (p. 15)

In class discussion, students also agreed that the learning method one personally feels comfortable to use may not be the most effective one for achieving a specific learning goal. They felt that they should be open to experimenting with new methods, and should monitor and evaluate regularly.

Metacognitive Strategies

The journal invited students to describe how they were planning, monitoring and evaluating the learning process, to enhance their awareness of the metacognitive strategies (see O'Malley, Chamot, Stewner-Manzanares, Küpper & Russo, 1985; Rubin, 1981, 1987; Seliger, 1984; Wenden, 1987). Planning and monitoring of out-of-class preparation was frequently evident in journal entries:

> "Everything went to plan. It gets easier every day I listen again to the tape."
>
> "I listened to the tape and read the dialogue a lot this week."

Long-term planning was also observed. These students were determined to get used to the sound of the language before using the script:

> "I plan to listen to the whole tape. Then, block by block, perhaps, listening to the tape, copying the intonation without looking at the script, and as it begins to sound more familiar, I will start reading from the script. . . ."

> "I hope to improve my fluency and intonation by 1) listening to the tape as often as possible, 2) listening to small sections and copying the intonation and 3) listening to small sections and being able to read along."

Monitoring and evaluating the method against the outcome was evident in the following journal entries:

> "Constant repetition, and writing lines out worked OK. However, taping my speaking may work better. Try that this week."
>
> "Sometimes I keep stopping after every line and other times I will go right through the section which helps me get an idea of the flow of the lines."
>
> Other conscious planning included dividing the task into manageable chunks:
>
> "I think I would probably go over my lines in the blocks that they are in, rather going through the whole play every time."
>
> "Long lines are difficult to memorize–it helps to break them up into smaller parts."

Those students separated the immediate task (learning correct pronunciation and intonation of a phrase) from a larger task (learning a whole sentence) and directed their attention to what they considered an achievable goal.

Paying Attention to Detail

Throughout the semester, students' skills in paying attention to detail were challenged. Two assigned tasks gave them specific opportunities to listen to their projection very carefully. The first of these was to record their lines and compare their tape to the model tape. All students perceived a difference. However, students varied in their ability to analyse the details. Some students heard only that there was something different, but could not identify the difference:

> "I'm not sure what it is. I can hear these differences. But I can not pin it down."
>
> "It's difficult to pin it down–because I don't think I can hear it very well."
>
> "I don't think I sound Japanese."
>
> "Mine is far less 'native' sounding!"
>
> "My accent is fairly Australian. . . ."

Some commented on the intonation:

> "I can't really pin it down. But as I mentioned above, my intonation is not quite the same as the taped version."

> "Yes, my accent is still fairly Australian and I have trouble following the ups and downs of intonation."
>
> "Yes, my recorded production sounds terrible. The intonation is not nearly as good or as clear as the tape."

One detail that many students had problems with was the intonation pattern of a "yes/no" question in Japanese. The English intonation pattern for a "yes/no" question was strongly imposed on the Japanese. It was very hard to avoid raising the pitch prior to the question particle ***ka***:

> "I also tend to raise my voice at the end of a question, instead of dropping it before *ka*."

Through careful study of the Japanese intonation patterns, one student discovered her habit in using rising intonation even for declarative sentences:

> "It was also hard not to raise the intonation at the end of a normal [declarative] sentence as we tend to do in English."

Young Australians tend to raise the pitch at the end of declarative sentences as if to say "Are you listening to me?"

Sometimes students were able to pinpoint and articulate the difference between their pronunciation and the native model:

> "I think my main problems are the devoiced vowels and nasalised /g/."

Students' words intuitively describing a certain discovery in their journals were brought into the classroom for elaboration, to help impart a more technical understanding of the perceived phenomena. The use of words such as "exaggerated" and "bounce" reflected the students' perception of the imposition of English intonation pattern into Japanese:

> "I think I'm trying too hard with my intonation, because some parts of what I've said sounds exaggerated."
>
> "I need to pronounce words with less 'bounce'–make them along the one level."

Another example was drawn from the next entry:

> "Mine is far less 'native sounding'! I can't consciously avoid pronouncing each and every syllable."

What this student meant was he was having a problem pronouncing vowels with even length. The class agreed that this was because he was imposing an English intonation pattern onto Japanese, making some vowels longer than others. The class discussed how English stress patterns affect the characteristics of syllables, resulting in the difference in vowel qualities. For example, in English, a stress-timed language, vowels in the stressed syllables are longer, louder and higher-pitched compared to weakened vowels in unstressed syllables. This does not happen in Japanese, where the intonation pattern is characterised mainly by pitch differences.

The use of journals thus constantly helped students connect the identification of specific phenomena they experienced to the wider perspective of problems experienced in foreign language learning.

In another task, students were encouraged to try to imagine how the lines should sound before they said them, and to compare that with what they heard or what they delivered, as some musicians are said to do. Most students said that they could imagine how the lines should sound. Some said that the imagined sounds were those of the native speakers which they heard on the tape, or heard spoken by the class teacher:

> "My imagined production is more likely an aural reproduction of a native speaker's speech."
>
> "In my head I can hear exactly how the tape sounded and try to repeat it. I don't know why it doesn't come out right sometimes, but I can hear the difference in my voice from what is in my head."
>
> "It is easier to think of the correct intonation in your head because you can remember what has been said by the teacher and this is in your memory."
>
> "I imagine how the teachers would pronounce the line and what their voices sound like."

Generally students perceived the gap between imagined sounds and their attempts:

> "Yes, (what I say) differs, because when it comes out, it is a lot slower."
>
> "The difficulty for me is knowing, and getting right when to change pitch."
>
> "No, no. I don't know. I think that what I imagined is right. However, my English intonation interferes when I speak."
>
> "It is very difficult to actualise what you mentally planned because what you think in your head never sounds like what you say. The imagined production does not really sound like what I actually say, I don't think. However it is very difficult to analyse this."

Many comments were made pointing out the difference in the speed of speech:

> "The tape is also said at a much faster speed and there are few breaks in the sentences."

> "I can not speak at the same speed as the people on the tape. This is very difficult."
>
> "I speak a lot slower than the tape."

Some students thought that the difference resulted from a lack of physical readiness to produce certain sounds and mimic intonation:

> "Since my tongue is not used to speaking with Japanese pronunciation and intonation, I might need to take some time to get used to them. That may be the reason why my actual production does not follow my mental production."
>
> "I imagined much better than the actual delivery. My tongue is very undexterous."
>
> "Getting my mouth (lips, tongue) to form the right shapes–this affects my pronunciation–sometimes I can clearly hear it in my head. But as it is so fast, . . . my mouth lags behind."
>
> "It is more difficult to say something as fast as you think it, especially for some particular sounds I am not used to, like the /r/ and sounds of the /k/ when devoiced."

Students and teachers accepted this lag between recognition and production ability as a part of the learning process. However, knowing that they could produce a much better quality projection by immediately repeating after the teacher in the classroom gave them confidence that the gap could be narrowed with persistent practice.

The Conflict Between the Substantial Memorization Task and the Accurate Learning of Pronunciation and Intonation

Although students' awareness of authentic pronunciation and intonation was gradually increasing, this was not necessarily reflected in the performance of many students in the first drama test, which was held in Week 8. Only a handful of students demonstrated convincingly good pronunciation and intonation in the test. The main reason attributed to this was that the assessment criteria also included memorization. Although students' memorization exceeded expectations, their pronunciation and intonation in general were not so accurate. Some students commented in the Week 8 journal after the test that they could not concentrate both on remembering their lines accurately and also on delivering with authentic pronunciation and intonation:

> "I did not think it was a very good test for our intonation ability as I had to concentrate too much on just memorising the lines rather than how to say them properly."

The teachers felt at this stage that the effort needed to approximate authentic pronunciation and intonation may well be in conflict with the effort to memorize the lines. This led the teachers to seek students' opinions as to whether acquiring native-like pronunciation and intonation conflicted with memorization. The Week 10 journal addressed this issue, and the responses were divided into two groups. Some said the two did not conflict:

> "There is no conflict if the lines are memorized by listening to the tape rather than just reading the script."
>
> "To fluently and authentically deliver my lines I have to know both the lines and intonation. If I don't know the rhythm and intonation, then I don't know my lines. For me I learn the intonation like a song, then within that structure, I have to fit sounds in. If I leave out a sound (a word or syllable) it sounds and feels wrong."
>
> "I don't think there is a conflict between delivering memorized lines and pronunciation and intonation. I think they're related, and when I was memorising my lines, I tried to perfect the intonation and pronunciation at the same time as memorising my lines. It was much easier that way."

Others said that the two conflicted:

> "Because I have to concentrate so much on remembering my lines I find it very hard to make any lines sound as fluent or as authentic in pronunciation as I would like."
>
> "My problem is that I have memorized the Romanised version of my lines [which the student produced for himself] and therefore my pronunciation and intonation may not be fluent because I will be delivering the lines from memory of the Romanised script."

From this it is clear that there was a difference perceived by students in the correlation between mastery of pronunciation and intonation and the effort put into memorization. Memorization strategy was important: those who said memorization did not conflict with pronunciation and intonation incorporated from the beginning the material they had to memorize with how that material sounded during out-of-class preparation. Students who said that they conflicted relied heavily on visual memorization, especially in the process of getting ready for the first test. Those students who did well in memorization but did not get good marks on pronunciation and intonation proposed in the same journal that they would concentrate on perfecting the prosodic quality of the memorized lines. However, some students commented that although they were prepared to work hard on correcting pronunciation and intonation for the final performance, it would be difficult, now that they had memorized the lines with their own interpretation of the sounds:

> "I find that whatever I learn first tends to stay with me. Therefore, if there is no effort, uncorrected mistakes remain - I think it takes more concentration to correct a mistake that's planted in my lines than to learn a new line."

The teachers' assessment of the accuracy of pronunciation and intonation in the first test led them to evaluate the appropriateness of the weight placed on memorization. They felt that students who sacrificed accurate pronunciation and intonation in the interest of memorization would not succeed in later correcting inaccurately learned pronunciation and intonation. This prediction was later found to be correct. Memorizing lines was a large task. Some sentences were complex and long, as authenticity required. Students also had to learn parts of others' lines so that they knew when it was their turn to speak. Listening to the tape alone did not give sufficient stimulus for learning lines. Students needed to use visual clues as well, and some students relied too much on the most comfortable input, the script.

The Final Outcome

For a beginners' course, the final (assessed) performances by most groups were of a high standard as productions. However, even those students who did well in the first test in pronunciation and intonation did not do as well in the final performance. This seemed to be due to their nervousness at performing in front of the audience, being assessed, and needing to pay attention to all the tasks associated with the performance:

> "The only trouble is when I am stressed or nervous I revert back to my original pronunciation even though I know the learned way as soon I finish my lines."
>
> "I think to have our pronunciation and intonation examined, we should have been examined separately because it was difficult to coordinate acting and lines as well as pronunciation. My pronunciation is much better when I just say all my lines."

This made the teachers wonder whether assessing pronunciation and intonation during the final performance was an appropriate method of assessment. I felt that although drama and performance were useful for learning native-like pronunciation and intonation, performance should not be exclusively used to assess pronunciation and intonation outcomes.

Students' Evaluation

Students experienced a great deal of satisfaction learning through the drama project, and a lot of positive comments were made in the post-project journal relating ~~to~~ perceived improvement in pronunciation and intonation:

> "I think [I] was all right before, but am much better now! Before, I didn't really know the Japanese intonation—only what I thought might be the Japanese intonation."
>
> "Improved a lot! I would be interested to listen to the first tape we made to see how much I improved."
>
> "As I had no experience of Japanese before, . . . my pronunciation and intonation has been moulded from the course. Drama has given me more confidence in Japanese. Yes, my Japanese has improved."

Some students commented that although they saw improvement, it was not perfect:

> "Much better, but still not very good."
>
> "Better intonation but my pronunciation does not seem a lot better. Some improvement. A long way to go."

It seems that, most importantly, this exercise enhanced an awareness of the importance of pronunciation and intonation, although there was a great variety in the degree of mastery of native-like pronunciation and intonation.

> "I think that since my pronunciation and intonation have improved, I feel that I have achieved something through this drama project, that I'm one step closer to being really fluent in Japanese."
>
> "Because there are certain Japanese phrases engraved in my head, those form the basis of what I can now hear as correct and incorrect intonation patterns."
>
> "I now know to make my intonation go up for questions and down for statements and so will be able to hear that when other people are speaking. Listening to the tape repeatedly helped me to understand the overall sound of sentences."

Some students stated that they gained confidence in speaking in Japanese:

> "It didn't hurt at all. My pronunciation and intonation did improve a bit. It was a very long learning process and I'm glad I've reached the end and been quite successful. I think it's given me greater confidence to speak Japanese in the future."

> "It's improved my confidence in speaking in Japanese, and given me a better understanding of how the language should be spoken."

Further research is needed to investigate whether or not or to what extent learning outcomes from the drama class can be transferred into the other aspects of language competence, such as listening ability and spontaneous speech, and also whether there are any long-term benefits which will be reflected in further progress by students.

Conclusion

Teachers felt that there were many areas for further improvement in the design of the drama project and the inquiry into whether a drama component can serve as an appropriate medium for learning native-like pronunciation and intonation. The length of the scenarios could have been shorter to lessen the overall amount of memorization for each student, thus avoiding conflict with accurate learning of pronunciation and intonation. Rather than working on a long play for the whole semester, two or more shorter pieces could have been practiced. The length of sentences could have been better controlled in the writing of the script. Although some students enjoyed dealing with longer sentences, these gave other students trouble. The assessment methods used also need revision. Although student progress towards the approximation of native-like pronunciation and intonation was evident in class rehearsals and through reading of the diaries, the formal grading based only on the two performances did not accurately reflect progress. Nervousness and the effort of memorization affected the results. The learning of pronunciation and intonation could have been made part of a continuing assessment process, based on regular performance in class. Alternatively, the group could have made a recording of the performance. Although it was not discussed in this study, working in groups created a problem when, as often happened, groups were missing a person or two in rehearsals. Each of these items presents major issues that need to be explored further.

However, above all, teachers and students felt strongly that student awareness of the learning process towards accurate pronunciation and intonation was greatly enhanced and that the use of reflective journals contributed most significantly to this process. Journals served not only as a tool to promote conscious learning by the students but proved beneficial in many other ways: they served as a rich mine of information from which to generate active class discussion and appropriate explanation and instruction; assisted teachers in making minor adjustments to the program; gave opportunities for every student to express themselves, especially beneficial for shy students or those who might not have expressed thoughts and feelings otherwise; fostered an open and supportive relationship between teachers and students during the learning experience; enhanced teachers' readiness and their ability to identify learning issues, specific or broad, by paying careful attention to students' words, which assisted in promoting successful teaching. Above all, reflective journals helped all participants develop a sense that both students and teachers were co-experimenters, co-observers and co-learners.

Harumi Moore holds an MA in Applied Linguistics in Japanese from the Australian National University. She is interested in the contrastive linguistics of Japanese and English and applied linguistics. She has taught at the Japan Centre at the ANU and various other institutions. She is currently studying towards a doctoral degree.

Note

1. Students at the Japan Centre include a number who have lived in Japan for up to one year. Those who had lived in Japan were placed in more advanced units from the beginning and did not take part in this drama project.

References

Allwright, D. (1983). Classroom-centered research on language teaching and learning: A brief historical overview. *TESOL Quarterly, 17*, 191-204.

Allwright, D., & Bailey, K.M. (1991). *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.

Asher, J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 8*, 334-341.

Bailey, K.M. (1983). Competitiveness and anxiety in adult second language acquisition. In H.W. Seliger & M.H. Long (Eds.), *Classroom-oriented research in second language acquisition* (pp. 67-103). Rowley, MA: Newbury House.

Bailey, K.M. (1990). The use of diary studies in teacher education programs. In J. C. Richards & D. Nunan (Eds.), *Second language teacher education* (pp. 215-226). Cambridge: Cambridge University Press.

Bogdan, R.C., & Biklan, S.K. (1982). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon.

Boud, D. (1986). Toward student responsibility for learning. In D. Boud, R. Keogh & D. Walker (Eds.), *Reflection: Turning experience into learning* (pp. 21-37). London: Kogan Page.

Brown, A. (Ed.) (1992). *Approaches to pronunciation*. London: Macmillan.

Busch, D. (1982). Introversion-extroversion and the EFL proficiency of Japanese students. *Language Learning, 32,* 109-132.

Chaudron, C. (1988). *Second language classroom research: Research on teaching and learning.* New York: Cambridge University Press.

Dickinson, L. (1987). *Self-instruction in language learning.* Cambridge: Cambridge University Press.

Guiora, A., Beit-Hallahmi, B., Brannon, R., Dull, C., & Scovel, T. (1972). The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry, 13,* 421- 428.

Holden, S. (1981). *Drama in language teaching.* Harlow, Essex: Longman.

Jones, M.R. (1985). Rapid learning, language acquisition, and the critical age question: The effect of a silent period on accent in adult second language learners. Unpublished doctoral dissertation, The University of Texas.

Krashen, S., Long, M., & Scarcella, R. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly, 13,* 573-582.

Kreeft-Peyton, J., & Reed, L. (1990). *Dialogue journal writing with nonnative speakers of English.* Alexandria, VA: TESOL.

Lamendella, J. (1977). General principles of neurofunctional organization and their manifestations in primary and non-primary language acquisition. *Language Learning, 27,* 155-196.

Lennenberg, E. (1967). *Biological foundations of language.* New York: John Wiley.

Maley, A., & Duff, A. (1982). *Drama techniques in language learning.* Cambridge: Cambridge University Press.

Matsumoto, K. (1987). Diary studies of second language acquisition: A critical overview. *JALT Journal, 9*(1), 17-34.

Morley, J. (Ed.) (1994). *Pronunciation pedagogy and theory: New views, new directions.* Washington DC: TESOL.

Neufeld, G. (1978). On the acquisition of prosodic and articulatory features in adult language learning. *Canadian Modern Language Review, 34,* 163-174.

O'Malley, J., Chamot, A., Stewner-Manzanares, G., Küpper, L., & Russo, R. (1985). Learning strategies used by beginning and intermediate ESL students. *Language Learning, 35,* 21-46.

O'Malley, J., & Chamot, A. (1990). *Learning strategies in second language acquisition.* Cambridge: Cambridge University Press.

Odlin, T. (1989). *Language transfer.* Cambridge: Cambridge University Press.

Olson, L., & Samuels, S. (1973). The relationship between age and accuracy of foreign language pronunciation. *Journal of Educational Research, 66,* 263-267.

Oxford, R. (1990). *Language learning strategies.* New York: Newbury House.

Oyama, S. (1976). A sensitive period in the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research, 5,* 261-285.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language acquisition. *Language Learning, 30,* 449-72. Reprinted in S.

Krashen, R. Scarcella, & M. Long (Eds.)(1982), *Child-adult differences in second language acquisition* (pp. 52-63). Rowley, MA: Newbury House.

Paulston, C. (1978). Biculturalism: Some reflections and speculations. *TESOL Quarterly, 12*, 369-380.

Peyton, J.K. (1990). *Students and teachers writing together: Perspectives on journal writing*. Alexandra, VI: TESOL.

Purcell, E., & Suter, R. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning, 30*, 271-287.

Richards, J. C., & Lockhard, C. (1994). *Reflective teaching in second language classrooms*. Cambridge: Cambridge University Press.

Rubin, J. (1975). What the 'good language learner' can teach us. *TESOL Quarterly, 9*, 41-51.

Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics, 11*, 117-131.

Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history and typology. In J. Rubin & A. Wenden (Eds.), *Learner strategies in language learning* (pp. 15 - 30). Englewood Cliffs, NJ: Prentice-Hall International.

Rubin, J., & Wenden, A. (Eds.). (1987). *Learner strategies in language learning*. Englewood Cliffs, NJ: Prentice-Hall International.

Schumann, F. (1980). Diary of a language learner: a further analysis. In S. D. Krashen & R. C. Scarcella (Eds.), *Issues in second language research: Selected papers of the Los Angeles second language acquisition research forum* (pp. 51-57). Rowley, MA: Newbury House.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House/Harper and Row.

Seliger, H.W. (1978). Implications of a multiple critical periods hypothesis for second language learning. In W. Ritchie (Ed.), *Second language acquisition research* (pp. 11-19). New York: Academic Press.

Seliger, H.W. (1984). Processing universals in second language acquisition. In F. Eckman, L. Bell, & D. Nelson (Eds.), *Universals of second language acquisition* (pp. 36-47). Rowley, MA: Newbury House.

Seliger, H.W., Krashen, S., & Ladefoged, P. (1975). Maturational constraints in the acquisition of second language accent. *Language Sciences, 36*, 20-22.

Snow, C., & Hoefnagel-Hohle. (1977). Age differences in pronunciation of foreign sounds. *Language and Speech, 20*(4), 357-365.

Staton, J. (1987). The power of responding in dialogue journals. In T. Fulwiler (Ed.), *The journal book* (pp. 47-63). Portsmouth, NH: Boynton/Cook.

Staton, J., Shuy, R. W., Peyton, J. K., & Reed, L. (Eds.) (1988). *Dialogue journal communication: Classroom, linguistic, social and cognitive views*. Norwood, NJ: Ablex.

Stern, S. (1980). Drama in second language learning from a psycholinguistic perspective. *Language Learning, 30*, 77-97.

Suter, R. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning, 26*, 233-253.

Tench, P. (1981). *Pronunciation skills*. London: Macmillan.

Wenden, A. (1985). Learner strategies. *TESOL Newsletter, 19*(5), 4-7.

Wenden, A. (1987). Metacognition: An expanded view of the cognitive abilities of L2 learners. *Language Learning, 37,* 573-597.

Wessels, C. (1987). *Resource books for teachers: Drama.* Oxford: Oxford University Press.

Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation.* Englewood Cliffs, NJ: Prentice-Hall.

(Received November 12, 1996; revised March 17, 1997)

Appendix: Sample journal questions

(Pre-project journal, Week 2)

Could you please tell us your overall feeing about the task lying ahead of you?

Do you feel embarrassed about blindly mimicking Japanese sound and intonation? If so, what do you think you can do to overcome such a feeling?

What kind of existing knowledge or strategies do you think you will use to help you mimic the pronunciation and intonation?

How do you feel about memorizing your lines in Japanese? How are you going to manage this task? Does it seem difficult? What makes it difficult?

How do you feel about working in a group? What kinds of benefits and difficulties can you foresee?

(Asked each week)

How are you feeling about your learning experience in drama in general?

(Week 3)

What kind of cues were you paying attention to most in class when you were practicing the lines? (1) listening to the teacher's lines and copying them? (2) looking at the Japanese script? (3) combination of above or any other method? Do you think that the method you used worked for you?

How do you plan to improve fluency outside the class: how much practice will you do at one time; what method(s) are you planning to use?

(Week 4)

Did the out-of-class learning go as you planned? What made it difficult? What helped? How many hours did you spend in practice last week?

What method(s) did you use? Do you think it (they) worked?

How are you checking your pronunciation and intonation?

(Week 7)

How do you feel about learning lines which are beyond your grammatical knowledge?

How does your learning experience through the drama project relate to the learning of other components of Spoken Japanese 1?

What did you think of the drama test 1? What do you have to do individually from now on? What do you have to do as a group from now on?

(Week 8)

For those who felt earlier on that not knowing exactly what you are saying interferes with the memorization process, how do you feel about it now after the first test?

(Week 10)

Do you think that there is a conflict between your concentration on delivering the lines from memory and concentration on how fluently you will deliver them with authentic pronunciation and intonation? Do the two efforts conflict with each other or can they be incorporated successfully? How do you propose to do this task successfully?

(Post-project journal)

In order to improve Japanese pronunciation and intonation, do you think the drama project provided you with an appropriate environment and help?

If you think you have improved your pronunciation and intonation, what did you discover and experience from going through such a process, and how do you think and feel about having achieved that goal?

What kind of long-term benefit will this experience have on you?

オーストラリアにおける観光業用の日本語コースのデザインと実践

鷲生ふさ子・舛見蘇弘美・トムソン木下千尋

ニューサウスウエールズ大学

海外からオーストラリアにやって来る観光客の２２％は日本人である。オーストラリアの観光産業にとって、日本語は非常に重要な言語といえる。本稿では、シドニー、ニューサウスウエールズ大学における観光業日本語コースのデザイン過程と実践、そして、その評価を教育学、社会言語学の理論に照らしながら、考察する。

コースはKnowles (1980)の社会言語学の理論枠組みを使い、言語能力、社会言語能力、社会文化能力、そして、自律学習の領域で学習目標を立てた。コースは地域社会のリソースを活用し、専門家の講義や、実地体験、インタビューなどを通して、学習を計った。評価の結果、教材の選択、言語能力に関する学習項目と、それ以外の学習項目のバランスに問題点が残った。

Japanese tourists comprise 22% of all in-coming tourists to Australia. The Japanese language is of extreme importance to the Australian tourism and hospitality industry. This paper discusses course design and delivery of "Hospitality Japanese" at the University of New South Wales. The course was designed using theoretical frameworks and recent research findings in education and sociolinguistics.

The course design followed the procedure proposed by Knowles (1980). It includes needs assessment in five different areas such as needs of the hospitality industry, needs of the hospitality industry workers, knees of the university and its language program. Recent research findings in sociolinguistics in regard to Australian speakers of Japanese in the hospitality industry also gave significant contribution to the needs assessment. Instructional objectives were written in accordance with the framework of Neustupny (1985) covering three areas of linguistic, sociolinguistic and sociocultural competencies, and an additional area of learner autonomy.

The course was successfully delivered using local community resources, such as guest speakers and field trips. However, the evaluation of the course showed two areas that needed improvement. One area was its teaching materials selection. The other was on the balance between the linguistic objectives and non-linguistic objectives.

JALT Journal, Vol. 19, No. 2, November, 1997

シドニーにあるニューサウスウェールズ大学、商業経済学部に所属するアジア・ビジネス言語学科（School of Asian Business and Language Studies、以下、当学科）では、オーストラリアの観光業界における日本語の重要性に着目し、既存のビジネス日本語コースに続き、1995年後期に観光業に関する日本語のコース（Hospitality Japanese、以下、観光業日本語）を新設した。ビジネス日本語をはじめとする特別な目的と特別な学生層を想定した専門外国語コース（Languages for specific purposes）には、一般的な総合外国語コースよりも、さらに厳しく綿密なコース・デザインが要求される。教育の分野におけるコース・デザインの理論、また、社会言語学の理論と最近の研究成果を実際のコース・デザインに活かし、実践したのがこの観光業日本語のコースである。本稿は、この観光業用日本語コースのコース・デザインの過程と実践、そして、その評価を、教育学、社会言語学の理論に照らしながら、考察するものである。

コース・デザイン

ある特定の教科のコースは、そのニーズの特殊性や、学習環境での条件、制約等が十分に考慮された上で、デザインされなければならない。Knowles (1980)は、コース・デザインの過程として、下記を提唱している。

1）コースを運営する組織内の環境を整える。
2）ニーズ分析をする。
3）コースの目標設定をする。
4）コースをデザインする。
5）コースを運営する。
6）コースを評価する。
7）1に戻り、このサイクルを繰り返す。

以下、本稿では、Knowlesの段階に沿って、観光業日本語コースのデザインの過程を追っていきたいと思う。

環境整備

コースの成功のためには、その運営に適した環境の設定が必須である。観光業日本語のコースの運営のために、次のような準備が行われた。

1）担当教官とシニア教官がチームを組んで、コース・デザインに臨む体勢を作る。
2）様々な器材を使ってのマルチ・メディア教育が可能な教室を確保する。
3）履修予定者を商学部の学生に限る。（学部によって入学規準が違うので一学部に限ることによって、学生の能力の幅を限定することになる。）
4）履修者を日本語学習初級中期の者に限る。（日本語レベルの幅を限る。）
5）学習者数の上限を20とする。

BEST COPY AVAILABLE

ニーズ分析

ニーズ分析は、学生のニーズに限らず、コースを運営する組織のニーズ、コースの関わる地域社会のニーズなど、多岐に渡るのが理想的である。このコースでは、以下のような分野でのニーズ分析を行った。

1）オーストラリア、特にシドニー観光業界の状況調査。
2）観光産業就業者の日本語のニーズ調査。
3）先行研究に表れる観光産業における日本語使用の問題点の調査。
4）コースを運営するアジア・ビジネス言語学科の方針と、その所属団体であるニューサウスウエールズ大学のニーズの考察。
5）コースを履修する学生のニーズ調査。

観光業界の状況

オーストラリアでは日本語は政府の指定した優先言語の一つとなっている(National Asian Languages and Culture Working Group, 1994)。日本がオーストラリアの最大の輸出相手国であること、世界各国からの観光客の中で、日本からの観光客の数が隣国のニュージーランドからの観光客を抜いて最多であること（同上）などを考えると、オーストラリアにとって日本語は最も重要なビジネス言語だと言っても過言ではない。

観光業界だけを取ってみても、オーストラリアへの日本人観光客の総数は1994年には69万人、1995年には74万人で（Australian Bureau of Statistics, 1996）、海外からの観光客総数の22％を占めている(Australia-Japan Society of Queensland, 1995)。また、観光客一人当たりのオーストラリアにおける出費も94年度には＄1645と、他国からの観光客に比べてひときわ高く（Bureau of Tourism Research, 1996)、オーストラリア経済にとって、日本人観光客を満足させるということの重要性がよくわかる。

日本人観光客が現地採用の日本人の店員にいらだちを感じる、また、オーストラリア人の店員にサービスを受けたいと思っているという調査結果(ATC, 1990)や、また、オーストラリア滞在中にオーストラリア人と一度も言葉をかわしたことがない日本人観光客が多く、なんらかの形で、オーストラリア人と直接のコミュニケーションを体験したいと言う素朴な願望も報告されている(Blackman, et. al., 1994)。また、オーストラリア人日本語話者の雇用先として、観光業界が非常に有望だと言う報告もあり(Marriott, et. al. 1994; Thomson, 1996)、有能なオーストラリア人日本語話者を観光業界内に配置することは有利といえる。

つまり、オーストラリアの観光業界には、日本語話者、特に有能なオーストラリア人日本語話者の養成というニーズがある。

観光業界で働く日本語話者のニーズ

続いて、観光業にたずさわるオーストラリア人スタッフと日本人旅行者の間で、実際場面の相互交渉の中から起こる諸問題を調査した。当学科で日本語学習中の学生の中から、日本語既習時間350時間以上の学習者で、実際に日本人観光客を相手に商業

的に応対の経験を有している者を選び（40人）、アンケート調査を実施し、18人から回答を得た。アンケートでは、　日本人客との接客に関する社員研修の有無と内容、接客方法、接客上のトラブルとその処理方法、　四技能に分けた日本語の使用状況について質問をした。

この調査結果の中から、観光業日本語のコース・デザインに関与する部分をまとめてみると下記のようなことが言える。

1）日本人客との接客のために日本語の表現レベルでの指導、たとえば、販売基本表現の「いらっしゃいませ」「何をお探しでしょうか」等や、挨拶表現は教えられているが、それだけでは不足で、日本人の習慣や好み、身振り言語に関する知識、及び「お客さんがいつも正しい」という商業文化の理解も必要である。

2）上記のような技能、知識面だけではなく、常に笑顔を保つ努力とか、あくまでも友好的に物事を処理するといった、就業態度に関することも重要である。

3）日本語の四技能の中では「話す」、「聴く」が等分に重視され、「読む」、「書く」技能の比重はわずかである。

先行研究からの示唆

観光業日本語関連の先行研究のほとんどは実際の観光産業の日本語使用の豪日接触場面の問題をオーディオ・テープまたはビデオで観察記録し、Neustupnyの理論枠組(Neustupny, 1985)を用いて分析した実態研究である。その中から、今回の観光業日本語のデザインに関連の深いものを挙げてみる。

一連の先行研究の先駆けとなったFukuda(1989)では、土産物屋でのインターアクションで、オーストラリア人の店員は日本人店員に比べると、日本人観光客にものを買わせるという実質行動ができないと報告している。つまり、客と店員の間でコミュニケーション行動は起こっているが、本来の目的であるはずの商売が成り立っていないのである。Marriott and Yamada (1991)も、オーストラリア人店員は基本的な言語能力は有しているが、客商売における社会言語学的なポライトネスを表現する能力に欠けていると報告している。Usuki (1993)は、日本の客商売についての一般的知識が大きく欠如しているために問題が起こるとしている。たとえば、客に満足の行くような商品情報、客と店員の目上、目下関係、海外旅行における買い物の占める位置、買い物量等についての知識がないことが問題となる。

オーストラリア人店員の言語行動の中で日本人評価者から最も否定的な評価を受ける項目として、前述の適切なセールス技術の欠如のほかに、客とやり取りをするときの適切な改まり語彙の量とバラエティーの不足、狭義の敬語のうち特に主体尊敬語の不足、イントネーションやピッチが乱暴で丁寧さに欠けることも報告されている(Usuki, 1993)。Enomoto (1993)は、オーストラリア人ツアーガイドの行動のビデオ分析から、日本人に最も否定的な評価を受けるのは発話行為（謝罪、賛辞、依頼、手助けの提供）における不適切さであり、非言語的要素（微笑、お辞儀、姿勢、身ごなし）も同程度の否定的評価を受けるが、狭義の敬語に於ては、誤用が意志の伝達に関わる

ものでないかぎり、上記ほどには強い否定評価を受けないと報告している。

このような先行研究の結果は、上記の観光業就業者の提示したニーズとつながるところがあり、前述のNeustupnyの枠組みである言語能力、社会言語能力、社会文化能力の三領域に渡っている。言語能力においては、狭義の敬語のうち主体尊敬語の強化が、社会言語能力については、ネウストプニー（1982）の点火、セッティング、参加者、バラエティー、内容、形、媒体、操作、運用のそれぞれのルールの体系的な指導が、また、社会文化能力については日豪の接客業の相違、日本人観光客、商品やサービスに関する知識の指導が必要である。

大学と学科のニーズ

上記は、大学を取り巻く地域社会を中心としたニーズであるが、大学における観光業日本語は観光業界内の社員研修や、民間の日本語学校の接客コースとは自ずからニーズが違ってくる。特に大学教育がエリート教育であるオーストラリアでは、大学卒業生が接客業に就くことは現実的でない。将来、観光業界で管理職に就く学生達の養成と考えるべきである。

また、このコースを直接に担当する当学科には日本語プログラム全体の目標が設定してあり、その目標にも添ったコースとしていくべきである。目標は下記のようなものである。

1）文法的、機能的、かつ文化的にも適切な日本語で、日本人とコミュニケーションができるような学習者を育てる。
2）異文化コミュニケーションへの理解力を持ち、さらに、日本人や日本に対して積極的な理解を持つような学習者を育てる。
3）自主的に自律学習が続けられるような学習者を育てる。

ニーズ分析

上記の様々な情報を加味し、総合的なニーズというものを考えてみた。大学出の管理職を想定したコースであり、日々の接客業務は考えられないものの、接客フロアで問題が起こった場合最終的な問題処理を任されるのは管理職の日本語話者である(Thomson, 1995b)ことから、接客に必要な言語能力を管理職が有していることは必要不可欠である。また、管理職が接客フロアの問題点を把握することの重要性も考えられる。従って、前述の言語、社会言語、社会文化能力の三領域に、自律学習の技能を加え、言語能力に関しては会話に重点をおいたコースを作るというのがこのニーズ分析の結果である。

学生のニーズ

上記までのニーズ分析はコースのデザインの前に得ることができた情報で行ったが、このコースを実際に履修する学生のニーズは、コースが始まった第一週目に初めて得ることができた。従って、コース運営上この情報を適用するという形での利用となったが、関連情報を下記に挙げておく。

履修した学生は13人で、その中で観光業関係のアルバイト経験のあるものは2名に

過ぎず、アルバイト先として、また、卒業後の就職先として観光業界を考えている履修者が9名(69％）いた。コースへの要望としては、言語、文化理解、観光事業関係の知識の三領域に要望が見られ、コース開始前のニーズの三領域とほぼ合致している。

学習目標の設定

ニーズ分析より下記のような学習目標を設定した。

1）言語能力目標：学習者は、日本人観光客との接客に必要な、場面に合った適切な敬語（特に主体尊敬語）能力をつける。ハワイ大学での"Travel Industry Management" コースでの教材開発研究でも言及されているように (Hijirida and Iwamura, 1986)、観光業日本語の場合一般的な日本語コースで考慮される四技能習得重視ではなく、会話を中心としたコミュニケーション技術習得に焦点をあてる。

2）社会言語能力目標：学習者は、日本人観光客に適切に対応するために、狭義の日本語のみならず、その社会言語的ルールを理解し、その知識を駆使して客のメッセージを読み取り、自己表現できる技術を習得する。

3）社会文化能力目標：学習者は、商業的立場から、日本人観光客への応対をより深く理解するために、日本の接客業についての知識を得、日豪の観光産業の現状を知り、日本人観光客がオーストラリア旅行に期待するもの、実際に扱う商品、サービスに関する知識を得る。

4）自律学習技能目標：学習者は、コースの活動を通して、コース終了後も観光業日本語の学習が続けていけるような技能を養成すると共に、観光業関連のネットワークを作る。

コース・デザイン

コースには、上記の大きなコース目標のほかに、学習内容、細かい学習目標、学習のためのリソース、学習過程、評価といった要素が含まれる(Thomson, 1995a)。

学習内容と学習目標

Beresford (1994)の来豪旅行者の言語調査によると、旅行者と日本語話者の接点は、航空会社、機内、空港、免税、土産物店、ホテル、レストラン、旅行会社、観光地、ツーリスト・アトラクションなどである。学習者に実際に観察、参加させることができ、社会言語、社会文化の能力を高めるのに有益で、しかも初級中期の言語力でも効果のあげることができる場面ということから、免税・土産物店での「買い物」「ホテル」「レストラン」の三種を選び、このコースの学習テーマとした。それぞれのテーマ毎に細かい学習目標を設定しているが、ここでは「買い物」を例にとって紹介する。

1）敬語使用による接客日本語の習得
・敬語法を学習する

BEST COPY AVAILABLE

・場面で使用される語彙を習得する
・免税、みやげ物店の接客場面での会話を分析する

2）社会言語能力がともなうコミュニケーション技術の習得
・お辞儀が正しくできる
・話し掛けるタイミングがわかる

3）社会文化的知識を付ける
・贈答習慣（おみやげ、せんべつ、おかえし）を知る
・ブランド志向文化を理解する

4）免税、みやげ物店でのフィールド・ワークで自律学習の力をつける
・現場の日本人とのインターアクションから、自己モニター力をつける

教材とリソース

教材は、コースの構想に完璧に合致した既成品が望めないことから、多少近い内容を持った教科書と、外部資料などから考案した副教材作ってそれを併用することにした。

さて日本語教師が学生のニーズにあわせたコース・デザインをするとき、限界を感じるのが言語以外での教師側の専門的知識の欠如である。また物理的制約の大きい教室学習のみでは十分な学習ができない。そこで、外部リソースを活用することが必要になってくる。地域社会のリソース活用については、トムソン（1996）に詳しいが、人的、物的、社会的、情報サービスなど、現地のリソースをできるだけコース・デザインに活用することが、よい学習結果につながる。

この観光業日本語コースでは現地のリソースを以下のように活用した。

1）オーストラリア観光産業の日本人観光客の動向に詳しい専門家を呼んで講演を聞く。
2）日本人対象のみやげ物屋、免税品店を訪問し、店員の接客方法等を観察する。
3）ホテルを訪問し、日本人ホテルマネージャーにインタビューし、問題点を探る。
4）懐石料理の調理師を呼び、試食しながらマナーを習う。
5）観光スポットに行き、日本人観光客にインタビューをする。

学習過程

このコースでは、上記のように、伝統的な教室活動に留まらず、様々な学習過程を通して学習目標の達成をはかった。会話は練習するだけでなく、学生に談話分析をさせることによって、敬語使用に関する理解を深め、また、練習した会話は、教室に現地の日本人を招き、学生に実際に使わせてみることで、習得の度合いを高めた。社会言語能力、社会文化能力の開発は、英語の補助教材やビデオ教材を使い、また、専門家を招いたり、訪ねたりしての実際のインターアクションを通して学習した。また、フィールド・ワークを通じ、現場でオーストラリア人スタッフと日本人観光客のイン

ターアクションを観察し、自ら参加して、実地体験することからも学習した。フィールド・ワークは、依頼から最後のお礼に至るまで実際使用の中でコミュニケーションの学習を可能にすると同時に、学習者に自分で自分の行動（特に社会言語学的ルールの適用について）をモニターし、かつ、現在の問題を解決するには何をすることが可能かを考える習慣を付けさせることもできる。

評価

評価方法は、指導内容、指導目標を考慮し、出席（１５％）、聴解テスト２回（１０％）、非言語を含めたコミュニケーションテスト２回（20％）、テーマ別プロジェクト（フィールド・ワーク、報告書）（35％）、プロジェクトの発表（20％）とした。

コースの運営

以上のようなコースを42時間で運営した。運営に当たって、履修学生のニーズや、学生との交流から、また、外部リソース利用の可否等の要因から臨機応変にデザインを変更していった。

コースの評価

コースの評価は、学習者の成績、学習者による評価をもとにこの筆者グループが考察をを行った。

学期末の成績が学習者の教科の習得と合致するとは言い切れないが、教師が設定したいくつかの評価規準の達成率という観点から、このコースの学習者の成績もコースの評価に示唆するものがある。学習者の成績の平均値は75点で、オーストラリアのシステムではDistinction（75-84、優秀）に当たり、最高成績のHigh Distinction (85-100)に続く成績である。

つまり、学習者はこのコースの評価規準に関して言えば、かなり高い度合いで目標達成をしたと言える。

当大学では、学期の終わりに履修学生による教科と担当教師についての２種類の評価が義務づけられている。ここでは教科評価のみにふれるが、学習者は、学習目標の明示や、評価方法、　物理的学習環境は適切であり、また、教室学習、専門家の講演、フィールド・ワークや実演講習は有益だったと評価した。指導の質は全体的にみて良いと評価し、 他の教科と比較して、この教科のレベルは適切といえるし、 この教科を他の学生に推薦するという結果が出ている。一方、教科書はあまり役に立たなかったという意見や、　学習目標と実際に教えられたことが一致していないという意見もあった。

学習者の成績、評価から受ける全体の印象としては、新企画のこのコースが総括的には成功であったものの、教科書の適性、そして学習目標と教科内容の一致に関する学生の不満は検討されなければならない。

まず、教科書の適性だが、主教材として言語関係と文化、コミュニケーション関係

の教科書を2冊学生に指定したが、豊富な補助教材と生教材、リソースの存在の中で、教科書の影が薄れ、また、42時間のコースでは使いきれない量であったため、せっかく買った教科書が余り役に立たなかったという不満が多かった。次回に向けて、指定教材の再選定が必要だ。

次に、学習目標と指導内容の不一致感だが、いくつかの原因が考えられる。実際には、指導内容が明示された学習目標を反映するようにコースが運営されたのだが、学習者が一般的に「日本語」のコースに期待するものと、このコースの教えたものの違いが、この不満感の原因ではないだろうか。「日本語」のクラス（特に大学の）というのは、文法のドリルをしたり、筆記試験があったりするものだという学習者の持ち込んでくる先入観が、実際のコースの学習目標を上回ってしまったということだ。言語、社会言語、社会文化、自律学習の四領域に費やされた時間のバランスもその原因の一つと言える。初級中期という日本語力のため、観光産業に関する知識の導入や、討論は英語に頼ることが多かったが、狭義の日本語習得のために費やされた学習時間が比較的少なかったために、十分な練習ができず、よって達成感に疑問があったと考えられる。つまり、言語面での活動の全体における比重、評価の割合の妥当性、使用言語、日本語と英語、の比重に検討の余地がある。

新しいサイクル

翌年のコース・デザインに向けて、今年度の成功要因と見られる四領域に渡る目標設定、外部リソースを活用したフィールド・ワークによる学習や専門家のセミナーは継続するものの、教材は新しく選択し直すこととした。また、日本語と英語の比重に関しては、履修資格を上げて、初級後半の日本語力のものを対象とするコースとすることで、日本語の比重を上げることを可能にする。また、学習者の期待感とコース目標のギャップに関しては、コースの初めに観光業日本語の持つ意義、本稿で挙げたようなニーズ分析の過程などを学生と一緒に検討することで、学習者の合意を得る試みを行う。

終わりに

本稿は、ニューサウスウェールズ大学で1995年後期に新教科として加えられた観光業用日本語のコース・デザインの過程と実践について述べたものである。理論、研究と実践は往々にして別のものと捉えられがちであるが、ここでは教育学、社会言語学の理論と研究結果をもとに実際のコース・デザインと、コース運営を行った。この試みに力を得て、他の専門日本語コースのデザインも、理論、研究結果が反映されたものになっていってほしい。この観光業日本語のコースにも新しい理論、研究結果を随時取り入れ、さらによいものにしていきたい。

謝辞：本コースのために、多くの方々にお世話になりました。特に、オーストラリアの日本観光事業の専門家、ニューサウスウエールズ大学マーケティング学科のR・マーチ氏にこの場を借りてお礼を申し上げます。

參考文献

Australia-Japan Society of Queensland. (1995). *Special Edition Newsletter,* August 1995.

Australian Bureau of Statistics. (1996). *Table Catalogue 3401.1 Overseas Arrival and Departure*. Canberra: Australian Bureau of Statistics.

Australian Tourist Commission. (1992). *Tourism Pulse*, No. 7.

Beresford, D. (1994). Foreign language tourism markets. Issues Paper presented for Industry Action Workshop.

Blackman, S & Nadsen, G. (1994). The Degree of Satisfaction: Are Japanese Tourists Experiencing Real Australia? Unpublished project report for JAPN4001 Japanese Communication 4B, Asian Studies Unit, The University of New South Wales.

Bureau of Tourism Research. (1996). *International Visitor Survey Information 1996*. Canberra: Bureau of Tourism Research.

Enomoto, S. (1993). *Native Speaker Reactions to Non-native Speakers' Deviations from Politeness Norms in Japanese Tour Guiding Situation*. Unpublished M.A. Thesis, Monash University.

Fukuda, H. (1989). *Japanese Tourists in Melbourne: Problems in Shopping Situations*. Unpublished M.A. Preliminary Dissertation, Monash University.

Hijirida, K. & Iwamura, S. G. (1986). Language for travel industry managers: Focus on Japanese. *Foreign Language Annals, 19*(1), 19-24.

Knowles, M. (1980). *The Modern Practice of Adult Education*. Chicago: Follett Publishing Company.

Marriott, H.E. & Yamada, N. (1991). Japanese discourse in tourism shopping situations. A paper presented at Seventh Biennial Conference, Japanese Studies Association of Australia.

Marriott, H., Neustupny, J.V. & Spence-Brown, R. (1994). Unlocking Australia's Language Potential: Profile of 9 key languages in *Australia*, Vol. 7. Canberra: National Languages and Literacy Institute of Australia.

National Asian Languages and Cultures Working Group (1994). *Asian Languages and Australia's Economic Future: A Report Prepared for the Council of Australian Governments on a Proposed National Asian Languages/Studies Strategy for Australian Schools*. Queensland: Queensland Printing.

Neustupny, J.V. (1985). Language norms in Australian-Japanese contact situations. In Clyne, M. (ed.). *Pacific Linguistics*, C-92, Australia, Meeting Place of Languages. (pp.161-170). Canberra: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.

Thomson, C.K. (1995a). A learner-centred foreign language curriculum for a large, diverse group of students? Yes, it is possible. *Australian Review of Applied Linguistics, 18*(12), 119-130.

Thomson, C.K. (1995b). Japanese language needs and communication problems of Australian non-native speakers of Japanese in the Sydney Japanese speaking business community. A paper presented at the 1995 Japanese Studies Association of Australia Conference, Brisbane, July.

Thomson, C.K. (1996). Employment and language ability of Australian non-native speakers of Japanese: A Sydney case study. *Japanese Studies Bulletin, 16*(1),15-34.

Usuki, K. (1993). *Deviation of non-native Speakers and Evaluative Behaviors of Native Speakers in Tourism Shopping Situations*. Unpublished M.A. thesis, Monash University.

トムソン木下千尋（1996).「日本語教育におけるリソース再考」日本語教育学会第五回研究発表論文

ネウストプニー J.V. (1982).『外国人とのコミュニケーション』岩波新書

（1996年9月10日受け）

Perspectives

Empathy and Teacher Development

John B. Kemp
Gakushuin and Seikei Universities

This article suggests that empathy plays an important role in a cross-cultural setting, particularly when foreign teachers work within the Japanese educational system. Three areas are identified, namely involuntary, emergent, and consciously attempted empathy. These can be thought of as separate categories or ideas which frequently overlap. The use of balanced and informed empathy can often facilitate the resolution of classroom communication breakdowns and can make a significant contribution to ongoing teacher development.

この記事は、異文化接触場面、特に外国人の教師が日本の教育制度の中で仕事をするときには、共感が重要な役割を果たすことを示す。共感に関して、三つの特徴が挙げられる。それは、主体的であること、進化すること、意識的に試みられることである。これらは、しばしば重なりあう別々のカテゴリーまたは概念とみなされる。バランスのとれた、知識にもとづいた根拠のある共感は、教室内でのコミュニケーションの障害を克服するのに役立ち、教師の成長に多大な貢献をする。

In his article on "Training Development and Teacher Education" in *The Language Teacher*, Underhill (1990, p. 3) defines teacher development as, "being essentially concerned with the effects that the teacher herself has on the learners and on the learning atmosphere of the class, as distinct from the effect of her techniques and materials." Elsewhere Underhill writes (1992, p. 71) of, "teachers . . . continually in the process of actualizing their own expectations, of becoming the unique and best teacher it is in them to be."

Quoting the work of Carl Rogers, Underhill emphasizes the importance of empathy, acceptance and authenticity, these being, "characteristics of good teachers which could be further developed in any teacher who had the commitment to do so" (Underhill 1990, p. 3).

JALT Journal, Vol. 19, No. 2, November, 1997

The following discussion focuses on empathy, the first of these characteristics, as a way of promoting teacher development in the Japanese university or college context.

Three Areas of Empathy

The term empathy is used here to mean identifying oneself with the feelings, perceptions and thoughts of another. Three types are discussed in this report: *involuntary empathy*, *emergent empathy*, and *empathy* resulting from a conscious effort to imaginatively and knowledgeably take the role of the other, hereafter called *consciously attempted empathy*. As constructs suggested by this writer, they are set out to enhance awareness of what many teachers already know: that by "feeling oneself into the situation of the other person," development as both a teacher and a person can be promoted.

Trying to enhance awareness of the familiar by presenting it so that it can be freshly perceived is not readily achieved through a step by step process, as might be the case when learning a concrete skill. Similarly it is not akin to the type of understanding used in mathematics, where, by division and subsequent arrangement of the constituent parts in a particular order, certain propositions can be derived from the preceding ones. Furthermore, the personality of the teacher might make one type of empathy more meaningful than another. Underlying this point is the view that human subjectivity both influences and is influenced by the theoretical and professional concerns of teaching. Thus, either within the empathy types themselves or externally with reference to teacher-student relationships and teacher approaches, there is paradox and often contradiction.

The following sections examine how different types of empathy might function in the language classroom.

Involuntary Empathy

Involuntary empathy is a foundation for all positive human relationships, including those between teachers and students. Where there are acceptable levels of classroom harmony, it is likely that the expectations of both learner and teacher are sufficiently realized for learning to take place. However, there can be communication breakdowns and, in the case of a foreign teacher in Japan, such breakdowns might be due to a clash of expectations derived from quite different cultures. Although it is true that resolution can be facilitated through appropriate use of empathy, it may also occur without any conscious attempt to empathize with the situation of the students. Thus, difficulties such as those set out by Stapleton (1995, pp. 13-16), "Why don't my students have opinions?" "Why are they so unquestioning?" "Why don't they talk and discuss?" and "Why are they so willing to memorize?" can be solved by the teacher through a process of stock taking, self-interrogation and discussion with informed colleagues. Here, involuntary empathy may operate but may not be directly recognized as assisting in the resolution process.

Emergent Empathy

A conscious attempt to empathize with the students' situation provides an additional route to such resolution. However for consciously attempted empathy to be more than of very limited value, the teacher-as-empathist has to make progress in two areas. She must gain knowledge of the students' culture and, equally important, of herself as well. The term *emergent empathy* is used in this discussion to describe the development of such knowledge. The empathy process has no end point, but is rather in a continual state of becoming.

Andic, quoting the last journal entry of Simone Weil, the French social philosopher, pacifist and mystic who died in England in 1943, writes,

> The most important part of teaching is to *teach* what it is to *know* . . . Nurses who really know read the pain of their patients as human suffering, feel it as their own, and therefore act, to relieve it, according to their suffering. (Andic's italics) (1993, p. 145)

In this passage, the nurse "reads" the pain of her patient through the operation of emergent empathy. It should be noted that "read" has special connotations in Weil's writing, in part because of her view that self-centredness causes us to interpret people, events and nature incorrectly. Here, the nurse must not only empathize with the patient but also must know how to alleviate the pain. Similarly, during conscious attempts to empathize with the students' situation, the foreign language teacher must be aware that self-centeredness distorts a correct understanding of others and that, however imperfectly, this understanding must be informed by knowledge of the students' world view, beliefs and values.

Knowledge about Japanese customs, civilization and achievements can be assimilated without direct involvement with Japanese people, for example through reading and study. Thus, knowing about the role of Confucianism in Japanese education will give insight into the questions Stapleton (1995) has raised. However being knowledgeable is not the same as being what Bennett (1996, p. 6) calls *interculturally competent*. This comes through not only knowing about the historical and developmental background of a country, but also through communication with its people. This communication must be informed in two ways. The foreign teacher has to be aware that some of the norms of her own culture might well be ethnocentric or objectionable to people of another culture. One way of determining which aspects of the teacher's culture might present problems to EFL learners is to recognize those cultural patterns which have been viewed negatively by informed outsiders. For example, Barnlund (1989, pp. 186-187) refers to some problematic aspects of American culture and quotes a Japanese person as saying, "the American love of freedom has exceeded all reasonable bounds and threatens to degenerate into self-centeredness." Jones (1984, p. 74) identifies negative stereotypes of "typical" American and British people: "arrogant," "uncultured," "reserved" and "hypocritical." Teachers should be on their guard against classroom displays of qualities which are problematic for people of other cultures.

Secondly, there has to be some understanding of what beliefs and values are assigned importance by the local culture. In the case of Japan, Hioki (cited in Loveday, 1986, p. 100) suggests the following: seniority, politeness, communal responsibility and sensitivity to face, inner versus outer worlds, modesty and the abandonment of individual self for a more collective identification.

From a Kantian perspective, new concepts and new ideas only have meaning when they are related to pre-existing experience and knowledge. Thus the extent to which the new is accepted and adopted when a foreign teacher deals with Japanese students and the educational setting in general is related to her pre-existing knowledge, values and beliefs. Suggesting which concepts and ideas are appropriate for a foreign language teacher in Japan is outside the limits of this article. However the relationship between personal values and teaching must be recognized. In the words of Edge (1996, p. 10), "Because we are people-who-teach (indivisible the person from the teacher), our actions in teaching arise from the same sources as our other actions and express deeply held values" (Edge's parentheses).

Consciously Attempted Empathy

Samovar, Porter, and Jain (1981, pp. 203-210) outline a six-step procedure to develop empathy. They first stress the importance of self-knowledge on the part of the empathist both at the level of personal ethnocentrism and at the cultural level of being aware of the image portrayed to the rest of the world by the empathist's country of origin. They also emphasize the need to be sensitive to feedback and to defer conclusions, especially when dealing with unfamiliar world views, values, languages, and nonverbal codes.

Their six-step procedure is as follows: 1) differences should be assumed among individuals and cultures, as not all people see the world in the same way. 2) We should know ourselves. 3) We should then temporarily set aside our self-identity and then 4) imaginatively put ourselves in the other person's situation. This leads to 5) the empathetic experience. Thus, having made a conscious effort to divest ourselves of our customary outlook, persona and ethnocentricities, we make a serious attempt to "walk a mile in the other person's shoes" in order to see and experience the world from their point of view. This is followed by 6) the re-establishing of our former self. We become once again the people we were before the careful reflection involved in the empathetic experience. However if the experience has resulted in insight into the other person's situation, the "pre-empathetic self" will not be exactly the same as the "new self."

A Personal Note

In my early days of teaching Japanese university students, I probably over-empathized with the students' situation. I had little or no grasp of the checks and balances outlined in the preceding two sections.

This is the land, according to Lebra (1976), where individuality, "rests not on the imposition of one's will on the social environment but on the refusal to impose oneself on it." Thus, after years of, from a westerner's point of view, self-development taking second place to the clearing of examination hurdles, I got the impression that, in the freer atmosphere of the university, the students would prefer a "social relations" approach to English. Informal chats with colleagues seemed to confirm this; although looking back, I realize that I was probably seeking confirmation of what I had already half decided rather than being open to alternative opinions. Pairwork, information-gap exercises, the sharing of personal information, games, mingling activities, and occasionally reading a short article from an English language newspaper became the order of the day with large freshman classes.

In addition, it did not take long to realize that basic humor which the students could easily understand really seemed to lift the classes. I could readily identify with Shimizu's (1995, p. 5) comment that, after nine years of teaching in Japan, she still feels that students view her as "more of an entertainer than a teacher." However, in my case, I overdid the entertainment side of things. "Playing to the gallery" seemed to be an essential part of the lesson. To my way of thinking, I was putting into practice Holliday's (1994, p. 113) maxim: "Learning about the real world of a new culture is a two-way reflexive process. One learns about others through monitoring how they respond to one's own actions."

Jokes and "social relations" English animated students, whereas listening and grammar exercises, together with "serious" topics, such as discussions about environmental issues, did not. Looking at the teacher and the lesson through the eyes of the students, and as a result giving them what they seemed to want, also found echoes in an existential phase I had passed through in my own student days.

It is hard to pinpoint when I changed the apparently successful formula of social English with an entertainment ingredient, but many teachers pass through a similar evolution. Richards (1994, p. 403) quotes Floden and Huberman on the three seasons of a teacher's professional life: stabilization, stock taking or self-interrogation, and disengagement. Pennington (1995, p. 705) quotes Freeman's observation that the key ingredient to teacher change and long-term development is awareness, a point emphasized by Kemp (1995) when discussing ways of recognizing cultural schisms.

A more informed understanding of the students' outlook and expectations showed clear differences between what I had thought about their expectations of university life and what they actually desired. Certainly there were those who wanted a social relations focus both inside and outside the classroom. But there were also those who wanted to continue to study. With such students as a catalyst, it was possible to interest the rest of the class in a wider range of activities and subject areas.

Classes are quieter now. The entertainment side of things is still important, even if much more occasional. A lesson might well be spent at the interface of what the students know and do not know about the use of "will" and "be going to." If they seem sufficiently receptive, the next week might be spent on "green" issues while the third week might center round a pop music and fashion video clip, with an accompanying likes and dislikes work sheet.

Concluding Remarks

This article has pointed to empathetic awareness in the context of broader personal growth as a way of promoting teacher development. The areas of involuntary, emergent and consciously attempted empathy can be thought of both as vertically separated categories and as horizontally linked ideas which frequently overlap. Which particular aspect is of relevance to the teacher will depend on her outlook, values and beliefs, together with her strength of commitment to awareness and action.

Consciously attempting an orderly six-step empathy training exercise might be more appropriate for someone who tends towards "convergent thinking," in contrast to a less methodical, more intuitive endeavor to assimilate knowledge of the second culture and of the self, as set out under emergent empathy. Peer help as a practical way of expanding and reinforcing the strength of commitment to awareness and action can be gained by what Edge (1992) calls "cooperative development." He sets out a framework of structured activities, which could well include a focus on enhanced awareness of empathy, designed to show how colleagues, working together, can promote self-development.

Kramsch (1993, p. 3) suggests that teaching is a juggling act which needs an intuitive grasp of the situation together with, "personal judgment based on as broad and differentiated an understanding as possible about what is going on at that particular moment in the classroom." It is suggested that balanced empathy informing day to day teaching decisions can make a significant contribution to the breadth and depth of such personal judgment.

John B. Kemp teaches at Gakushuin and Seikei Universities.

References

Andic, M. (1993). Discernment and the imagination. In R.H. Bell, (Ed.). *Simone Weil's philosophy of culture* (pp. 116-149). Cambridge: Cambridge University Press.

Barnlund, D.C. (1989). *Communicative styles of Japanese and Americans: Images and realities*. Belmont, CA: Wadsworth.

Bennett, M. (1996). Intercultural communication in a multicultural society. *TESOL Matters, 6*, 6.

Edge, J. (1992). *Cooperative development*. Harlow, Essex: Longman.

Edge, J. (1996). Cross-cultural paradoxes in a profession of values. *TESOL Quarterly, 30*, 9-30.

Holliday, A. (1994). *Appropriate methodology and social context*. Cambridge: Cambridge University Press.

Jones, L. (1984). *Ideas*. Cambridge: Cambridge University Press.

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford: Oxford University press.

Lebra, T.S. (1976). *Japanese patterns of behavior*. Honolulu: University of Hawaii

Press.

Loveday, L. (1986). *Explorations in Japanese sociolinguistics*. Amsterdam & Philadelphia: Benjamins.

Pennington, M.C. (1995). The teacher change cycle. *TESOL Quarterly, 29*, 705-731.

Richards, K. (1994). From guessing what teachers think to finding out what teachers know: the need for a research agenda. *TESOL Quarterly, 28*, 401-404.

Samovar, L.A., Porter, R.E., & Jain, N.C. (1981). *Understanding intercultural communication*. Belmont: Wadsworth Publishing.

Shimizu, K., (1995). Japanese college student attitudes towards English teachers: a survey. *The Language Teacher, 19*(10), 5-8.

Stapleton, P. (1995). The role of Confucianism in Japanese education. *The Language Teacher, 19*(4) 13-16.

Underhill, A. (1990). Training, development and teacher education. *The Language Teacher, 14*(1), 3, 12.

Underhill, A. (1992). Process in humanistic education. *ELT Journal, 43*(4), 250-260.

(Received November 25, 1996; revised February 2, 1997)

Answer, Please Answer! A Perspective on Japanese University Students' Silent Response to Questions

Timothy J. Korst
Ryukyu University

Many EFL teachers in Japan have noticed that Japanese students seldom have the skills to answer questions appropriately. In response to questions, teachers often hear only silence. This paper examines the issue of a silent response to teacher questions in the EFL university classroom and presents three activities designed to introduce sociolinguistic skills into the communicative syllabus. Instead of avoiding this issue by ceasing to ask questions, it is suggested that teachers should attempt to address and remedy this silence directly.

日本で教える英語教師の多くは、日本の大学生が質問に適切に答えるスキルをもっていないことが多いのに気づいている。質問しても、沈黙に迎えられることもしばしばである。この論文は、大学の英語の授業での教師の質問に対する沈黙の反応を検討し、コミュニカティブなシラバスに社会言語学的スキルを含めるための三つの学習活動を紹介する。質問するのをやめることで問題を避けるのではなく、この沈黙を直接、解決する努力をすべきであるということが示唆される。

Japanese university students pose a special problem for many EFL teachers. Due to reasons such as the Japanese cultural and educational background and the concept of saving face (Ishii & Bruneau, 1991), these students often remain silent when asked a question in English. The effects of such behavior can be twofold. First, the pace of instruction may be slowed while the teacher waits for an answer. Second, because this behavior is not normal within the Western classroom context, foreign EFL teachers may become frustrated by the silence. Confronted with this problem, I suspect that teachers often deal with it by avoidance, by adjusting their teaching techniques to avoid eliciting such behavior. Indeed, Mutch (1995, p. 14) seems to recommend avoidance of questions in front of the class, noting that teachers can create a more "relaxed and intimate atmosphere" through, for example, pair work.

JALT Journal, Vol. 19, No. 2, November, 1997

However, silence in response to a question is seldom an appropriate response in English communication (see Savignon, 1983). Answering questions is a basic skill to be mastered in learning a foreign language. This paper recommends dealing directly with the issue. It presents three activities designed to solve the problem of a silent response to questions and to promote the development of sociolinguistic skills.

Research on Classroom Interaction

Although student silence in response to teacher questions in the classroom has not been the focus of much EFL classroom research (Larsen-Freeman & Long, 1991), Miller's 1995 investigation indicates that students usually remain silent out of fear of limitations in their English speaking ability (also see Ishii & Bruneau, 1991). Investigating students' attitudes towards misbehavior in the classroom, Ryan (1995) found that both Japanese and Australian university students rated silent responses as relatively minor infractions. On the other hand, considerable interaction research has focused on teacher talk and teacher questions (see Chaudron, 1988). Turn-taking in the classroom (Sinclair and Coulthard, 1975) and comprehensible input (Ellis, 1994) have also been important concerns for researchers. Moreover, error correction and teacher feedback are common topics in the literature (see reviews in Chaudron, 1988 and Ellis, 1994). Such research focuses indicate that teacher questioning and giving-feedback is a common activity. Therefore, when the teacher asks a question and the students do not respond, this indicates the need for promoting greater communicative competence (Savignon, 1983).

Research on wait-time in ESL classes suggests that students may be able to produce answers if teachers wait slightly longer than usual. Studies in the ESL setting (see Shrum and Tech, 1985) suggest that the mean wait-time for ESL students is under two seconds. However, Holley and King (1971) found that teachers of German who waited at least five seconds obtained an increase in student responses. White and Lightbown (1984) and Long, Brock, Crookes, Deicke, Potter, and Zhang (1984) recommend that teachers should therefore allow a longer wait-time. However, while this approach may be beneficial with some students, allowing a longer wait-time with Japanese students may only compound the problem since an answer may not be forthcoming regardless of the interval (Ishii & Bruneau, 1991).

Other research (see Sato 1981, Long and Sato 1983, and Tsui, 1987) has investigated the type of teacher-student interaction which occurs in ESL classrooms. Using constructs similar to Long and Sato's (1983) "display" and "referential" questions, Tsui (1987), compared "social" and "classroom discourse." Social discourse differs from classroom discourse in that more negotiation of meaning occurs. Tsui (1987. p. 337) argues that interlocutors bring with them a set of "assumed shared beliefs which are constantly tested against, revised or added onto in the course of the interaction." Kartunnen (1973, cited in Tsui 1987, p. 337) speaks of the "common ground" that is achieved among the interlocutors, while Tsui speaks of the "social convergence" which is achieved by social discourse.

In classroom discourse, on the other hand, less negotiation usually takes place. The roles of the interlocutors (here, the teacher and student) are more clearly defined. The teacher is the "primary knower" (Berry 1981, cited in Tsui 1987. p. 339) who asks questions and the one who stands in judgment of the student's response. Transfer of knowledge is assumed to be unidirectional, from the teacher to the student. The student's role is usually to answer the teacher's question; she in turn evaluates the student's answer. The following is an example of classroom discourse given by Tsui (1987, p. 339):

(A)	T: Who can tell me what the two kinds of verbs are?	I
	Ange	I?
S:	Verbs of action and verbs of being.	R
T:	Right	F

Here, the labels I, R, and F refer to the "initiating move"(I move), "responding move" (R move) and "follow-up move" (F move).

Surely this pattern is common in EFL classrooms as well. By asking questions of this sort, the teacher seeks feedback from the students. If the students answer correctly, the teacher has achieved her objective and then can proceed. But what happens when the students do not answer? Tsui notes that a refusal to answer would be out of order; inability to do so would very likely be negatively evaluated by the teacher. Yet this is a common occurrence in the Japanese EFL classroom, and the teacher is left with the problem of how to evaluate the silence.

Types of Questions in Japanese College EFL Classrooms

In content-based college EFL classes, some of the classroom questions asked by the teacher comprise social discourse. This may include questions which deal with the content of the course as well as questions which are more personal in nature. Examples of such questions include:

(B) How was your weekend?
How are you today?
What's new?

It is through such questions that the EFL teacher hopes to create an atmosphere conducive to language learning. Though these questions are brief, they are not meant rhetorically and, in asking them, the teacher briefly steps out of her role as teacher and allows the students to view her in a different light. The questions may be asked privately to an individual student or to the entire class. Theoretically (see Krashen, 1987), such questions are intended to lower the students' affective filters, thereby creating a more relaxed learning environment. If the teacher is a native English speaker, the questions also allow the students to interact socially and give them the opportunity to realize that English can be used outside of the parameters of the lesson.

Teachers may also ask questions about the content of the lesson. For example, in a content-based English conversation class discussing the concept of the Third World, the following dialogue may take place (T is the teacher and S is the student).

(C)	T: Have you ever been to the Third World?	I
	S: Yes.	R
	T: Where did you go?	I
	S: Thailand.	R
	T: Uh huh.	F

This dialogue may be seen as two exchanges, again using the labels I, R, and F to refer to the "initiating move" (I), "responding move" (R) and "follow-up move" (F). Both exchanges represent social discourse.

But what happens when the following type of exchange occurs?

(D)	T: Have you ever been to the Third World?	I
	S: Yes.	R
	T: Where did you go?	I
	S: Canada.	R
	T: Canada is not in the Third World.	F

Here, the student has not answered according to the teacher's expectations, so the teacher has corrected the student. Perhaps in this case we could classify the first I and R moves as being social discourse and the latter I, R, and F moves as being the more formal classroom discourse (McCarthy, 1991). It is not so clear from Tsui's account whether teacher evaluation alone denotes classroom discourse.

In both C and D above the student is shown as responding to the teacher's questions. However, more realistically, the conversation might go like this:

(E) T: Have you ever been to the Third World?
S: (silence)
T: (more slowly) Have you ever been to the Third World?
S: (confers with neighbor, but does not reply)
T: Do you understand me?
S: (silence)

The following section describes three activities for avoiding this type of response and, instead, promoting the development of student communicative competence in answering both social and classroom questions.

Activities for Overcoming Classroom Silence

The activities presented below are designed to remedy the problem of silence in response to questions, to train students in answering within an acceptable amount of time and to give them practice in asking for clarification.

The activities were performed during a first year university content-based EFL conversation course during the last three months of the Spring, 1995 semester. The course met for two 90-minute periods per week. Of the 28 students, 22 were women. Although no proficiency test was given, the students' English proficiency was considered to be at the low-intermediate level by the author. Because of the nature of Activity 1, it was performed only once. The other two activities were regularly performed for 30 minutes per class period.

The three activities followed the same general format. The students paired off by forming an inner and outer circle and partnered the student in the opposing circle. They performed the activity then rotated in opposite directions, enabling the students to repeat the activity with a new partner from the opposing circle. This is illustrated in Figure 1.

Setting up the circles

The teacher can easily arrange the students into two circles by having them stand in alternation–inside, outside, inside, etc. An unpartnered student can form a triad with two others. After the students have formed the circles and performed the activity with their first partner, the teacher should tell those on the outside of the circle to turn to their right and find their partner two positions along the circle. Changing two positions to find new partners gives the students a sense of progressing around the circle, and also prevents them from looking ahead to their next partner.

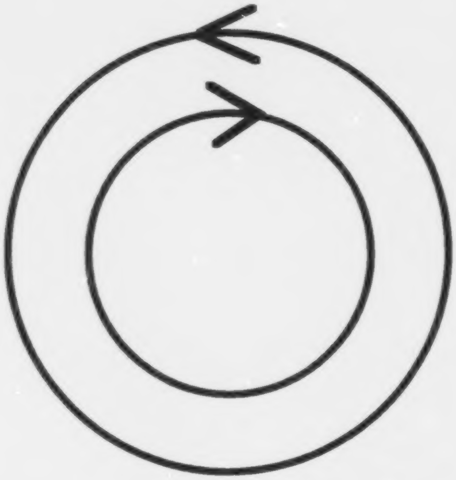


Figure 1: Students form inside and outside circles, pair up, and after doing the activity rotate to the right.

Activity 1

Purpose: To give the students the experience of asking someone a question and receiving no response. By performing this activity, the students may develop empathy with their teachers.

Giving Directions: After dividing the class into "inside" and "outside" groups, the teacher gives instructions to the two groups separately. While the inside students are still seated, the teacher tells them to put their heads down so that they cannot see. The teacher then writes the following instructions for the outsiders on the board.

1. Think of a question to ask your partner in the circle
2. Ask each partner the same question
3. Make sure you receive an answer from your partner

At this point the teacher may want to ask the students verbally if they understand the task. If students have questions, they can ask the teacher privately.

Once the teacher has given the instructions to the outsiders, she gives the following instructions to the insiders. At this time, the outside students have their heads down so that they cannot see the instructions. The teacher writes the following instructions on the board.

1. In the circle your partner will ask you a question
2. Don't answer it
3. Don't say anything

Once the instructions to both groups are clear, the students form their respective circles and perform the activity. Teachers can ask the students to change partners as often as they like.

Post-activity: The teacher tells the students to write a minimum of five sentences about this activity. Afterwards, the students discuss their ideas in groups and/or as a class. The teacher can write the students' comments on the board.

Activity 2

Purpose: To teach the students how to ask for clarification. Often the reason for silence is lack of linguistic knowledge, so the students must learn how to negotiate meaning and how repair a conversation when it breaks down. This activity allows the students to practice various ways of asking the questioner to repeat or explain the question.

Pre-activity: The teacher should explain what asking for clarification means. She can solicit various patterns from students and write them on the board. Some examples are:

(F) What did you say?
I don't understand.
Huh?
Could you explain that?
I'm sorry, what was that?
What do you mean?
Excuse me?/ Pardon me?

The teacher illustrates these techniques by asking the students questions that they cannot understand (i.e., by speaking too fast, using difficult vocabulary items and grammar structures, or mumbling), thereby soliciting the sentences given in Example F.

In performing this activity, the students must have a reason for asking for clarification. To insure this, the teacher can instruct students how to mumble (the other ways mentioned above may be too difficult for most students). The teacher should model mumbling a question, and have the students repeat as a class. Then the teacher should articulate the question clearly, and again have the students repeat. Once the students are adept mumblers, they should prepare questions to ask their partners and form their circles.

Activity: The students are paired up in inside and outside circles. Taking turns, the students ask each other questions while mumbling. Their partners then ask for clarification. The first student repeats the question clearly, and the second gives an answer. For example:

(G) A: (mumbles a question unintelligibly)
B: I'm sorry, what did you say?
A: What are you going to do on the weekend?
B: Go shopping.

The teacher can decide how many times the students should change partners.

Activity 3

Purpose: To teach the students to answer questions within an appropriate time period.

Pre-activity: The students prepare questions which they will ask their teacher. After a student asks a question, the teacher raises one hand and puts up a finger for each passing second, answering the question within five seconds. In this way the teacher demonstrates that it is appropriate to answer within a certain time limit. The teacher can also demonstrate that verbal responses such as fillers are also appropriate, but must be uttered within the same time limit. Some possibilities include:

(H) Uh . . .
Hmm . . . .
Well, let me see . . .

Such responses serve as a notice that the person questioned has understood that a question has been asked, but needs time to formulate an answer. The teacher can model these fillers and instruct the students to practice them through repetition.

Activity: The students prepare questions to ask their partners. If possible, the questions should be difficult, requiring some thinking time. The students again form inside and outside circles and ask their questions to their partners. After asking the question, the students should raise one hand and lift up one finger for each passing second. This serves as a visual reminder of the time limit for the second student, who should try to give a response within five seconds. After asking and answering questions, the students change partners.

Activities 2 and 3 can be combined. In circle formation, the students ask their partners questions, and receive one point each time they mumble (giving partner a chance to ask for clarification), ask for clarification and answer within five seconds. The students tally their own points, and whoever has the highest number wins. This game format can be integrated into course material whenever partner work is called for.

Student Reactions to the Activities

The students in my classes were able to carry out the three activities without much difficulty. However, minor problems arose. At the beginning of Activity 1, some students did not understand the directions written on the board. When I asked them verbally whether they had any questions, a few raised their hands, then approached me to ask their question in a whispered voice. Most often they asked whether they had to use the same question each time, to which I responded, "No."

In carrying out this activity, some of the "silent" partners answered with a nod of the head, or used other body language. In this case, I reminded the students to remember their instructions. Upon completion of the activity, the students wrote their reactions and shared these in groups. I wrote these responses on the board in two columns: outsiders and insiders. Almost all comments were negative. The outsiders, who asked the questions, described their negative feelings with comments like, "I feel very very lonery (sic). Why did they say nothing?" Another reported, "It is hard time and nurvous (sic) for me in today's class." Responses from the insiders, who kept silent, were also negative. One insider's response was, "I felt really awkward not answering the questions . . . . I think that is how foreigners feel sometimes." After the students realized that most feelings were negative, I spoke about my own feelings when students do not answer in class. "I feel the same way as you." I told the class. At this point I felt we reached a deeper level of understanding.

If most students felt uncomfortable doing Activity 1, many enjoyed learning how to mumble in Activity 2. This activity elicited much laughter and the students seemed to have no problem performing it. Activity 3, which focused on responding within a five-second time limit, proved a bit awkward. Quite a few students did not count to five. Perhaps some wanted to focus on the conversation, or possibly felt bad about imposing a limit on their partner. Likewise, the combination of Activities

2 and 3 into a game, with the tallying of points, proved difficult.

As a follow-up activity, the students wrote longer responses (about 300 words) in their journals. Some comments follow:

> As a result of our discussion we knew that the circle outsiders are foreigner and insiders are Japanese, or outsiders are teacher and insiders are students. . . . Because I was outsiders, I could understand teacher's feelings well.
>
> When I played that game, I understood foreigner's feelings. They feel lonely, irritate, nurvous, impatient. And we let them feel so. I feel very sorry about this fact.
>
> In the game, it was funny for me not to answer . . . . outsiders felt uncomfortable, irritated and bored. I didn't noticed that feelings before I heard that and I felt sorry for them. . . . What I learned through this game was how Tim felt in this class.

From the students' comments, it is evident that they understood the point of Activity 1. This activity proved successful in that it gave students insight into their behavior and the effects it has on others.

Questionnaire Results

After the students performed these activities for three months, they filled out the following questionnaire (See Table 1, below) in which the three activities were called "conversation games." Because this was a preliminary evaluation of teaching materials, no attempt was made to quantitatively evaluate improvements in the students' sociolinguistic skills. The students' response are given as the average number of points from a possible ten points for each statement.

Questions 1, 2, 3, 6, and 8 asked for the students' response to the activities. Here the mean response was above 5, suggesting that the students liked the activities although they found them difficult (Question 4). In particular, students indicated that they liked asking for clarification, perhaps suggesting that they may not have focused on this during prior English instruction.

Conclusions

Although I did not collect data, it is my impression that the students improved in their ability to give a verbal response to questions within an acceptable time limit. During the final oral test, most were able to ask for clarification in response to questions which I deliberately asked rapidly. During class, however, many still had problems. Even after

Table 1: Copy of Questionnaire and Results

Please answer the questions. CIRCLE A NUMBER. Thank you.

Results
x

1. I liked the conversation games.
NO 1 2 3 4 5 6 7 8 9 10 YES = 6.7

2. I think the conversation games are useful for learning English.
NO 1 2 3 4 5 6 7 8 9 10 YES = 7.6

3. At first, I did not like the conversation games. But now I like them more.
NO 1 2 3 4 5 6 7 8 9 10 YES = 7.0

4. The conversation games are difficult because I can't speak English very well.
NO 1 2 3 4 5 6 7 8 9 10 YES = 7.4

5. I don't like the conversation games because I have to talk with people I don't know.
NO 1 2 3 4 5 6 7 8 9 10 YES = 4.1

6. I like the conversation games because I can make friends with my classmates.
NO 1 2 3 4 5 6 7 8 9 10 YES = 8.0

7. In the conversation game, 5 seconds is too short to answer in.
NO 1 2 3 4 5 6 7 8 9 10 YES = 6.2

8. I liked practicing how to ask for clarification (What did you say?).
NO 1 2 3 4 5 6 7 8 9 10 YES = 7.9

9. I have trouble counting my points in the conversation game.
NO 1 2 3 4 5 6 7 8 9 10 YES = 5.2

several weeks of practice, many students' first option in attempting an answer was not to ask for clarification, but rather to consult with their neighbor. This indicates the need for continued focus on making a proper response.

It should also be noted that I received considerable emotional relief by engaging my students in these activities. Previously, student silence had an adverse effect, leaving me feeling frustrated and helpless. After these activities, however, I could view the students' silence with more objectivity and humor. When a student was silent in response to a question from me, I could say in a humorous voice, "I'm getting angry." Similarly, mumbling became a running joke.

Integration of these activities into the syllabus allowed me to view the students' silent behavior as something that we, as a class, could work on together. Addressing the problem directly eased the strain felt by both the teacher and students and through these activities I was able to assist my students to become more sociolinguistically competent in English.

Timothy J. Korst teaches at Ryukyu University.

References

Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Holley, F., & King, J. (1971). Imitation and correction in foreign language learning. *Modern Language Journal, 55*, 494-498.

Ishii, S. & Bruneau, T. (1991). Silence and silences in cross-cultural perspective: Japan and the United States. In Larry A. Samovar & Richard E. Porter (Eds.), *Intercultural communication* (pp.314-319). Belmont, CA: Wadsworth Publishing Company.

Krashen, S. D. (1987). *Principles and practice in second language acquisition*. New York: Prentice-Hall.

Larsen-Freeman, D. & Long, M. (Eds.), (1991). *An introduction to second language acquisition research*. New York: Longman.

Long, M., & Sato, C. (1983). Classroom foreigner talk discourse: forms and functions of teachers' questions. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in language learning* (pp. 268-285). Rowley, Mass: Newbury House.

Long, M., Brock, C., Crookes, G., Deike, C., Potter, L., & Zhang, S. (1984). *The effect of teachers' questioning patterns and wait-time on pupil participation in public high school classes in Hawaii for students of limited English proficiency*. (Technical Report No. 1). Honolulu, HI: University of Hawaii at Manoa, Center for Second Language Classroom Research, Social Science Research Institute.

McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.

Miller, T. (1995). Japanese learners' reactions to communicative English lessons. *JALT Journal, 17*(1), 31-49.

Mutch, B. M.. 1995. Motivation and cultural attitudes: Increasing language use in the classroom. *The Language Teacher, 19*(8), 14-15.

Ryan, S.M. (1995). Misbehavior in the university classroom? A cross-cultural survey of students' perceptions and expectations. *The Language Teacher, 19*(11), 13-16.

Sato, C. (1981). Ethnic styles in classroom discourse. In M. Hines & W. Rutherford (Eds.) *On TESOL '81* (pp. 11-24). Washington DC: TESOL.

Savignon, S. J. (1983). *Communicative competence: Theory and classroom practice*. Reading, MA: Addison-Wesley.

Sinclair J. and Coulthard, M. (1975). *Towards an analysis of discourse*. London: Oxford University Press.

Shrum, J. L., and Tech, V. (1985). Wait-time and the use of target or native languages. *Foreign Language Annals, 18*, 305-313.

Tsui, B. A. (1987). An analysis of different types of interaction in ESL classroom discourse. *IRAL, 25*(4), 336-353.

White, J. & Lightbown, P. (1984). Asking and answering in ESL classes. *Canadian Modern Language Review, 40*, 228-244.

(Received January 10, 1997; revised April 9, 1997)

A Poem in the Process: Haiku as an Alternative to Brainstorming

John Esposito
Kogei Women's Junior College

Studies of contrastive rhetoric now address not only product and process concerns, but also the complex rhetorical traditions which inform writing practice. By contrasting the rhetorical traditions of Japanese and English, it is possible to establish a point of convergence which can inform the teaching of writing to Japanese EFL students. This paper suggests that during the pre-writing stage of paragraph composition, haiku can be used as a complement to or substitute for brainstorming. A sample lesson is offered exploring some of the practical applications of this approach.

対照修辞学の研究は、書き上がりの産物と書く過程だけでなく、書き手の実践に影響を与える複雑な修辞学の伝統も対象とするようになった。日本語と英語の修辞学の伝統を対照することで、その類似性が発見できれば、外国語として英語を学ぶ日本の学生に英語で書くことを教えるために役にたつ。この論文は、パラグラフ・ライティングにおいて、書く前の段階で、俳句を使うことによって、文化的に適切な方法でブレインストーミングができることを示す。このアプローチの実践的応用を探究するために、授業のサンプルが紹介される。

> *We make out of the quarrel with others, rhetoric,*
> *but of the quarrel with ourselves, poetry.*
> *W.B. Yeats*

Over the past thirty years, theoretical approaches to second language writing instruction have centered primarily upon the issues of product and process (Raimes, 1991; Silva, 1990). This debate can essentially be construed as one of emphasis between the *what*, or the patterns, forms, and organization of texts, and the *how*, or the ways, uses, and functions of writing. At the center of this discussion has been contrastive rhetoric studies. Based on the Sapir-Whorf hypothesis that language influences thought, contrastive rhetoric studies began with Kaplan's seminal work concerning L2 student essays and the degree of

JALT Journal, *Vol. 19, No. 2, November, 1997*

negative transfer from a writer's L1 to L2. His main argument was that a student's native language influences second language acquisition, especially with respect to writing (1966). Kaplan's hypothesis, and, by extension, his definition of rhetoric, has been criticized as being too simplistic, reductionistic and ethnocentric (Liebman, 1992; Severino, 1993), yet, due to its intuitive appeal, it has influenced many second language teachers and researchers. Kaplan (Grabe & Kaplan, 1989; Kaplan, 1987), while subsequently qualifiying his initial theory as being overstated, insisted on the importance of teaching rhetorical forms, suggesting that they constitute a significant factor influencing L2 writers.

Overemphasis on linguistic accuracy and patterns led to a paradigm shift in second language writing research and pedagogy. The resulting process approach placed the writer's composing competence, as opposed to linguistic competence, at the center of attention (for an overview, see Krapels, 1990). Writing came to be seen as "a complex, recursive and creative process" (Silva, 1990, p.15), whereby the writer focuses on discovering meaning through the communicative purpose of a text. Students should therefore be provided with "ample time to write and rewrite, to learn that several drafts may be needed before intention and expression become one" (Zamel, 1982, p. 205). Proponents of a process approach also emphasize the fluid nature of a text, alternating between a pre-writing stage (concerned with brainstorming ideas, focusing and planning structures), a composing stage (where content, development and organization are addressed), and a post-writing or refining stage (characterized by drafting, editing and revising).

In recent years both content and audience awareness have also become important issues in second language writing instruction. In part this is a reaction to the primacy given the writer's cognitive needs in process methodologies. One major criticism of the process approach is that it neglects the sociocultural context in which writing takes place. Writers need to address the culture-specific forms and content acceptable and/or understandable to the reader. These necessities assume, of course, a preoccupation with form, or product. Horowitz (1986) states that writing without structure accomplishes little and that students should not be left to their own devices. An emphasis on form also has heuristic value as it motivates students to generate, invent and search for information (Coe, 1987; D'Angelo, 1980).

In spite of the apparent cyclical nature of L2 writing theory, characterized by adherence to rather narrow conceptualizations and prescriptions, there appears to be a growing awareness that the variables of text, writer, context and reader are not discrete, but are interrelated and should be addressed as such (Raimes, 1991; Silva, 1990). An eclectic approach to writing instruction, then, would include both product and process concerns, particularly rhetorical factors related to coherence and the cultural expectations of reader and writer (Connor, 1996; Grabe & Kaplan, 1989; Purves, 1988).

Concurrent with the shifting emphases of L2 writing research, contrastive rhetoric has evolved to consider not only the surface features and patterns of texts, but also the complex contextual dynamics underlying the writing process (Severino, 1993). Matalene (1985) defines this context as the relationship between culture, language and rhetoric. A corollary to this is the notion that cultural predispositions extend beyond the text to influence all areas of discourse (Leki, 1991; Strevens, 1987). Kaplan, for one, suggests that even logic is "a cultural artifact rather than an inherent capacity of the mind" (1990, p.10). Insofar as reasoning (i.e., the perceived relationships between phenomena) is culturally determined, it is limited to a writer's language. As a result, what may appear illogical to readers in one culture, is perfectly understandable to the readers of another (see Leki, 1992). Time also appears to be a cultural convention, which manifests itself in the arrangement of texts (Kaplan, 1990). Likewise, attitudes toward knowledge, including the approaches, strategies and aims of learning are closely tied to cultural norms and values (Ballard & Clanchy, 1991; Purves & Hawisher, 1990). These differences manifest themselves at the rhetorical level and often constitute barriers to communication (Strevens, 1987).

In orienting their readers to a topic, for instance, Scarcella found that second language learners' introduction strategies differed significantly in both quantity and quality from those of native English writers (1984). These behaviors seem to be reflective of particular education systems, where students "do indeed learn to become members of a rhetorical community" (Purves & Hawisher, 1990, p. 191). Through direct examination, contrastive rhetoric can be used to explicate culturally-informed aspects of writing such as implicitness and explicitness, clarity and coherence, unity, content, and other pragmatic concerns (Martin, 1991). According to Leki, contrastive rhetoric studies ultimately "concern themselves with the social construction of knowledge within discourse communities" (1991, p.135). In order to simplify their tasks, students first need to know what salient elements serve as the building blocks of this construction. Thus, there has been a strong call to make the learner aware of rhetorical differences through metalinguistic instruction (Carrell, 1987; Carson, 1992; Hinds, 1987; Hinkel, 1994; Jenkins & Hinds, 1987; Kaplan, 1990; Leki, 1992; Purves & Hawisher, 1990; Scarcella, 1984; Strevens, 1987).

It is my purpose here to show how this might be done with Japanese students of English. Although the field of contrastive rhetoric has focused primarily on negative transfer, there has been some support for use of the learner's L1 as an important resource (Cumming, 1989; Hinkel, 1994; Lay, 1982; Raimes, 1991). For example, Friedlander (1990) found that use of the L1 can assist students during the planning process as it facilitates information retrieval (see Carson, Carrell, Silberstein, Kroll & Kuehn, 1990). With this in mind, I would like to contrast what Liebman-Kleine refers to as "the rich views of rhetoric" (1986, p.8), which necessarily include the cultural traditions informing the whole writing process. First, I will try to indicate how the cultural and rhetorical traditions of Japanese and English influence present day usage. After highlighting two predominant and interrelated features of these disparate languages, I will suggest a point of convergence where they appear to complement each other. Finally, I will demonstrate, in the form of a lesson plan, how these insights might be exploited by both second/foreign language teachers and students of English.

A Contrastive Analysis of Japanese and English Rhetoric

Japanese Rhetoric

Two salient features of Japanese texts are indirectness and, consequently, the reader's responsibility to construct meaning. Historically, this might be traced to the Heian Era (794-1185), when waka (short 31-syllable poems) were exchanged by members of the nobility to communicate their love for each other (Tsujimura, 1987). Waka, which are still composed as tanka today, are characterized by their indirect and allusive wording. One vivid example of this is the way in which Heian poets referred to colors. The Japanese language originally had only four words to denote different colors. As a result, writers chose natural objects to evoke the myriad images of color they had in mind (Ooka, 1991). The Japanese propensity to be indirect might also be attributed to political factors such as the emphasis put on restraint during feudal times, and the need to conform, at least, outwardly, to the dictates of totalitarian regimes (Tsujimura, 1987).

Conformity, however, is not solely a consequence of sociopolitical exigencies, but is also a product of Confucianism. Ballard and Clanchy (1991) argue that due to the Confucian dictum that knowledge and truth are fixed and simply passed on from teachers to students, there is a strong aversion to argue or critically analyze with the intent to reach clear-cut conclusions (see also Hinds, 1982). For a Japanese student, there is thus "a willingness to tolerate ambiguity, even contradictions, to allow them to sit easily in tension within the same piece of writing" (Ballard & Clanchy, 1991, p.33). This seems to echo Hinds (1987), who notes the Japanese disinclination to explain or clarify when writing. In addition, because of the Buddhist and Confucian stress on maintaining harmony, "language is understood as a medium for expressing social cohesion, and not primarily as a medium for individual expression" (Carson, 1992, p.42).

Another factor influencing Japanese rhetoric is the nature of the Japanese writing system. Although writing classes are generally dispensed with by the end of junior high school, a substantial amount of time is expended on learning grammar and kanji (Chinese characters). In fact, Japanese students are compelled to learn four distinct written codes (five, counting Arabic numerals): 1850 kanji, which are the most formidable—ranging in complexity from 1-23 strokes that are written in a specified order—with most also having multiple readings (as many as nine depending on context); two phonetic syllabaries (hiragana and katakana) consisting of 46 basic symbols each; and the Roman alphabet. One result is that much time is spent on drill and memorization learning this intricate system. To compound this, Japanese words are not separated at the sentence level, thus compelling the reader to intuit the beginning and ends of words. This lack of clarity extends to the essay level as well, where transitions are usually not marked or attenuated (Hinds, 1987). One can plainly see how the language itself helps to foster ambiguity (see Clancy, 1986), thus placing cognitive demands upon the reader to recover meaning from a text.

Japanese writers expect that their readers' minds will work in similar ways to their own (Hinds, 1990). As Kaplan (1988) makes clear, when one is addressing culturally diverse groups of unknown readers, "the probability of shared universes of knowledge diminishes in direct proportion to the size of audience" (p.284). The potential audience for any text composed in Japanese is limited for the most part to the people living in Japan, the only country in the world where Japanese is used as a primary language. It comes as no surprise, then, that Japanese texts tend to be reader-responsible (Hinds, 1987). One by-product of a reader-responsible rhetoric is the value Japanese seem to place on expressive writing, that is, writing done about and for the self, at the expense of writing done with communicative intentions (e.g., persuasive or expository prose). As a consequence, Japanese readers expect that they will have to evaluate a writer's propositions on any given topic (Hinds, 1990). Indeed, it appears that Japanese writers are not very concerned with audience at all. They consider the beauty, or aesthetic aspect, of a text and engaging the reader's emotions to be good qualities of writing (Dennet, cited in Leki, 1992). Furthermore, writing is not thought of as a process of discovery, but comes only after thinking. Liebman (1992) found very little attention given to revising, most of which was limited to making sentence-level corrections (cf. Hinds, 1987). This lack of a heuristic aspect to writing is problematic as it precludes the whole argument concerning process as a recursive, dynamic search for meaning.

In Japan students learn to write through very little or no direct instruction (Mok, 1993). At the rhetorical level, writing is taught mainly by following formulas and through reading. When writing a business letter, for instance, Japanese depend to a great extent on choosing set expressions from style manuals (Jenkins & Hinds, 1987). Reading instruction includes encouraging the habit of reading between the lines while analyzing syntactic relationships (Carson, 1992). Essentially, learning to read and write in Japanese is learning individual words—words which are not generally used to convey ideas but for social functions. This word-boundedness also appears to carry over to the learning of English, where Yamada (1993) asserts that writing is limited to the sentence level as discourse and rhetorical organization are ignored (see Kobayashi, 1984).

The reading texts used as models are generally selected from Japan's long and distinguished literary canon (Liebman, 1992). The one rhetorical pattern that is predominant in traditional Japanese literature, indeed, in most of Japanese culture, is the *JO-HA-KYUU*. Ueda (1967) likens this pattern to the three movements of a western sonata (Exposition–Development–Recapitulation). In fact, this pattern is the basis for traditional Japanese music, where the *JO* is characterized by a quiet tone with a slow tempo, the *HA* incorporates a lighter mood and leisurely changes, and the *KYUU* increases the rhythm as well as the impact. This *JO-HA-KYUU* pattern is so pervasive that it is also found in traditional Japanese football, the Noh drama, renga (linked verse), the Kyougen (traditional Japanese comedy), and the Tea Ceremony. The fact that Japanese students learn to write primarily by reading would lead one to assume that the *JO-HA-KYUU* pattern must be internalized to some degree. Although Hinds (1982) claims that this in not the case at the compositional level, Mok (1993) feels that this organizational pattern probably forms the basis of Japanese writing practice.

English Rhetoric

The roots of English rhetoric can be properly traced to the models of ancient Greece and Rome. The first rhetoricians, the Sophists (circa 500

B.C.), were public speakers who argued on behalf of matters of sociopolitical importance. Their milieu was the public square and other places where people naturally congregated. They sought to provide their listeners with the necessary reasoning and arguments to make informed decisions (Saunders, 1970). Consequently, it was the listener's responsibility to form opinions through critical evaluation. This ability to evaluate the truth forms the basis of the Western rhetorical tradition. Whereas the Sophists emphasized the art of persuasive speaking, Plato sought to give preeminence to the search for truth (Hare, 1982). For Plato, contemplation, not argumentation, formed the basis of this search. Aristotle (see Roberts & Bywater, 1984) later reconciled these antithetical modes of inquiry by introducing a rhetorical system based on logic. Logical argumentation in the quest for truth thus came to be a defining factor of English rhetorical practice. Rhetorical standards such as clarity, coherence and linear progression also arose out of this oral tradition that produced the paragraph. Etymologically, the word is, in fact, derived from the Greek word paragraphos, which means a line in a dialogue showing a change in speakers.

The truth Aristotle had in mind was not catholic in nature but was open to debate. Whether by *ethos* (an appeal to the speaker's moral qualities), *pathos* (an appeal to the emotions of the audience), or *logos* (the logic of the subject matter), the way one argued was based upon the premise that ideas exist prior to and independent of language (Knoblauch & Brannon, 1984). This notion, in turn, also influenced form: a speech should be arranged linearly according to introduction, argument and counterargument, and summary. Consequently, the use of simple, ordinary words was, for Aristotle, the sine qua non of effective persuasion. He insisted that one must avoid ambiguity at all costs so as not to confuse or mislead the listener. Roman thinkers such as Cicero furthered this prescriptive view of rhetoric by focusing on the stylistic concerns of the speaker.

With the advent of movable type in the early 15th century, medieval and Renaissance scholars turned away from the spoken word to concentrate on writing (Lindemann, 1982). At this time there was apparently little difference between speech and its written representations; Hence, organization as an aid to coherence at the sentence level was of imminent concern. The confluence of Enlightenment thought, with its increasing faith in logical analysis and the scientific method of inquiry, and the emerging preoccupation with the written word, furthered the cause of short sentences and simple words (Bacon), expressing precisely the truth or falsity of propositions (Descartes). The length of the average sentence in English texts, in fact, was reduced in half from approximately the 16th to the 19th century (Rodgers, 1965). The concomitant increase in number of sentences, however, placed a greater burden on the reader. Rhetoricians thus shifted their attention to the paragraph, and the need to provide readers with comprehensible chunks of information.

Alexander Bain (1818-1903) was one such thinker. He regarded the paragraph as merely a big sentence (Rodgers, 1965). His six rules of the paragraph, which both mirror classical thinking and have influenced subsequent rhetoricians, was a prescriptive attempt at achieving coherence by avoiding ambiguity. His number one rule attests to his preoccupation with this issue: "The bearing of each sentence upon what preceded it shall be explicit and unmistakable" (Shearer, 1972, p.412). Bain's narrow prescriptions of what form a finished paragraph should take offered little insight into how a paragraph should be crafted, however. This formalistic view of knowledge eventually ran up against 20th century psychological interpretations that stressed the holistic and dynamic properties of cognition. The idea of learning through discovery, based upon Piaget's work on cognitive development, led to the belief that the focus of writing instruction should be on the process itself; thus bringing us to the source of our current strains in contrastive rhetoric.

While there continues to be much discussion over the definition of the paragraph (e.g., see Harris, 1990, for a study on the existence and/or placement of topic sentences), there is some consensus that modern English rhetoric tends to place a high value on clearly-reasoned, explicit, convincing prose. This is incumbent upon the writer to produce, to avoid miscommunication (Ballard & Clanchy, 1991; Hinds, 1987; Leki, 1992; Purves & Hawisher, 1990; Strevens, 1987). It would be unfair, however, to categorize all English or writing as exemplifying as having these values (Connor, 1996; Strevens, 1987) since rhetorical conventions are cultural as opposed to linguistic. Thus, the qualities of explicitness, clarity, and writer responsibility might best be viewed as tendencies of modern American English (Purves & Hawisher, 1990). However, the responsibility placed on the writer for avoiding miscommunication in modern English rhetoric is generally accepted.

A Point of Convergence

There are obviously great distinctions between the Japanese and English rhetorical traditions. The EFL writing teacher's primary concern is to make students aware of such differences while explicating such variables as coherence and the culture-specific expectations of both reader and writer. Given that the paragraph is an essential aid to coherence (Knoblauch & Brannon, 1984) and is a necessary convention of modern English prose, paragraph writing needs to be taught. One way might be to begin with the resources the learners bring to the classroom—in this case, the Japanese rhetorical tradition.

As mentioned earlier, Japanese texts tend to be ambiguous, which, in turn, places cognitive demands upon the reader. These qualities are best exemplified in the Japanese poetic form of haiku. The haiku was originally the opening sequence of the renga (linked verse). Renga were composed by several poets in a kind of lyrical interplay. After the first poet wrote the opening 5-7-5 part, known as the hokku, the next poet would add a two line 7-7 syllable conclusion. These two parts made an intelligible, independent whole (Sato, 1983). A point worth noting is that the haiku (hokku) was originally part of a larger pattern. And even after it began to be composed independently, it retained this fragmentary nature.

Haiku, like all poetry, is concerned essentially with experience. In three short lines it presents concrete images without any explanation. The symbolism of the central image or the relationships between images is suggested or hinted at. This is accomplished through such poetic devices as internal comparison, superimposition and juxtaposition. Because it is part of the Japanese rhetorical tradition, haiku is clearly reader responsible (Henderson, 1967). Japanese texts tend to assume a high degree of shared knowledge. It is the reader's job to make the connections and fill in the missing information in order to at once make sense of the poem and share in the poet's emotional response to the scene presented.

Regarding the form of haiku, some attempts have been made to impose a logical progression between the three parts of the poem. The haiku has in fact been likened to a sonata (Horiuchi, 1993; Ueda, 1967), a modern three-act play (Horiuchi, 1993), as well as a syllogism (Blyth, 1981). Horiuchi (1993) even claims that a haiku contains three ideas (thesis–antithesis–synthesis) that proceed in a rather linear manner. Of course, there are haiku that on the surface appear to be following a logical sequence, but such haiku are certainly the exception. Haiku simply do not have an internal logic per se. The connections between the parts of a haiku are accomplished through suggestion. If anything, a haiku is dependent on the reader to make a conscious connection between the parts. Thus, the *JO-HA-KYUU* is not a pattern generally present in the haiku, yet its influence can be detected in subtle ways.

Although there appears to be some merit in drawing analogies between the surface features of an English paragraph and the three parts of a haiku, it is suggested that the haiku is best exploited as an aid to the pre-writing process.

Haiku Writing as a Brainstorming Technique

One common device used to generate ideas in the pre-writing stage of paragraph composition is brainstorming. When writers brainstorm, they put down all the words or ideas that come to mind about a specific topic. The purpose of this unstructured probing according to Lefkowitz "is to help free your thoughts, break down mental blocks, and open your mind to other possible ways of looking at things" (1987, p.1). The ultimate goal of brainstorming is to flush out one's latent memory of all the items connected to a particular word or concept.

A haiku also seeks to codify in language one's unfettered thoughts. Ueda feels that the composing of haiku must be done in an instant "with no impure thought intervening in the process" (1967, p.159). Whereas brainstorming evolves from spontaneous connections between words, a haiku involves the poet's immediate response to the images or reality before one. Matsuo Basho (1644-1694), who is considered one of the greatest haiku poets, spoke of spontaneously expressing one's instantaneous perceptions. He also urged his students to speak their mind without wandering thoughts (Higginson, 1985). Therefore, a haiku is similar to brainstorming as they both emphasize writing without conscious intellectualization, without imposing one's subjective interpretation on the process.

A haiku is nonetheless an intellectual construction which depends upon descriptive accuracy as well as a heightened imagination. Masaoka Shiki (1867-1902), another master of haiku, felt strongly that words should reflect the image before one in order to create what he called "a sketch of life" (Beichman, 1986, p. 54 This sketch should be composed in the moment between perception and thought; in other words, before the brain becomes fully engaged. Shiki also taught his students to compose as many drafts as possible on a given subject when writing haiku. This seems to echo the writing teacher's admonition to write down everything that comes to mind. Another feature of haiku that appears to lend itself well to the pre-writing process is its universality, or ability to take in all the things of the natural world as subjects. Another feature of haiku that appears to lend itself well to the pre-writing process is its universality, or ability to take in all the things of the natural world as subjects.

A Sample Lesson

The following lesson was designed for Japanese first year university EFL students. It can be completed in two 90-minute class periods; the first day is devoted to writing the poem and the second day to converting the poem into a paragraph. The lesson can be adapted for other levels. In any event, encouragement will be needed, particularly on the first day as students tend to doubt their ability to compose anything significant.

Day One

1. Have students brainstorm words to describe each season. Since season words are an integral part of haiku, this is a good time to talk about the important role of nature in the poem. Brainstorming can be done by writing the seasons on the board one at a time and asking the students to write down the first word that comes to mind. A list should then be made either on the board or in the students' notebooks (see Higginson, 1985, for a list of about 600 season words and phrases in both Japanese and English).
2. Put the students into small groups. Then show them a photograph or print of a natural scene using pictures culled from magazines or old calendars, or ask the students to bring in their own pictures. Also, this could be a nice opportunity for an outdoor excursion. In any case, ask the students to once again write down words that describe the picture and their feelings associated with the scene.
3. Provide the students with an example of a haiku in order to explain its form (three lines), and content (description of a scene from nature using concrete language). Depending upon the students' level, the teacher might want to introduce more difficult aspects of the poem such as juxtaposition (the internal comparison of images), ellipsis (suggestion through understatement), or *kireji* (the cutting word or caesura). A good poem to use as a model would be Basho's famous haiku about the frog. Japanese know the poem by heart, and its surface features (depending upon the translation) compare favorably with the linear progression of most paragraphs. Also, there is an abundance of translations in English from which to choose (see Sato, 1983).
4. Ask the students to write the first and second lines of a haiku using their word lists from Parts 1 and 2. Urge the students to use clear,

simple language. Students can either work alone or in groups. I have found it beneficial to have students work in pairs for this part.

5. Have the students look at the picture or observe the scene again, considering the following two questions: (1) How does the picture/scene make you feel? (2) What does the picture/scene make you think about? The students need to have an emotional response to the scene because without one there really isn't a poem. This response should be conveyed, however subtly, to the reader. Then ask the students to write the third line of the poem by themselves.
6. At this point the students might want to share what they have written with others. As the class will most likely be working with the same photos or from similar physical stimuli, the students tend to be interested in what others have written. Nevertheless, this should be voluntary.

Day Two

The paragraph to be written here is expository or descriptive.

1. Present the students with a model of a paragraph based upon a haiku. You might want to compose your own paragraph from the haiku you used in Part 3 on Day One. or use an example written by a student. (See the appendix for some examples from my classes).
2. Discuss some of the differences between Japanese and English rhetoric, particularly how these differences pertain to the haiku they have written and the paragraph they are about to write. This discussion should be limited to how these differences help shape the two forms. It should be emphasized that the paragraph will be their explanation of the poem. This would also be a good time to review sentence development and paragraph organization.
3. Order is more important in the paragraph than the haiku, so the teacher might want to help students with this transition. The teacher could also show how haiku writing is similar to other pre-writing techniques, such as outlining or list-making.
4. Subsequent drafting and peer editing should provide students with enough opportunities to polish their work.
5. Once again, sharing is a natural way of bringing closure. Poetry in particular is made to be spoken, but class or department publications are also fun. Likewise, student-generated collages or haiku contests can elicit a lot of creativity.

Conclusion

Like all poetry, haiku is essentially ambiguous and suggests more than it states. What is suggested then becomes the basic idea for the paragraph. When turning a haiku into a paragraph, the writer clears away ambiguity through explanation. Thus, the paragraph becomes the writer's interpretation of the reality of the poem. This process involves the use of pre-writing and revising strategies as two vital components of good writing practice. Another plus for haiku is that their brevity forces a deeper, more disciplined approach to language (Higginson, 1985). Brevity requires the poet to leave out unnecessary grammatical words and connectives. These are often the parts of speech which give Japanese learners of English the most difficulty. Moreover, the sentence fragments or phrases that make up the poem will later be used in the paragraph. Finally, the success of writing a poem in a foreign language will undoubtedly have a positive effect on students' confidence (see Hirose & Sasaki, 1994).

The idea that poetry can be the basis of rhetorical instruction is not something new. In the West such a pedagogical approach can properly be dated to the first century AD, when the Roman rhetorician Quintilian spoke of the utility of turning poetry into prose. In those days literary language was seen as the foundation of good communication. This strong connection between poetry and rhetoric continued through the Renaissance; when poetic analysis informed writing pedagogy. In the first colleges of colonial America, students learned to write by reading Latin and Greek, a good percentage of which was poetry. Even in the earlier part of this century, literary texts were used as material for analysis in order to teach writing. Hence, the introduction of haiku to the process of paragraph writing might very well be seen as a cross-cultural variation of this tradition.

Acknowledgments

The author would like to thank Paul Hilderbrant and two anonymous JALT Journal *readers for the valuable suggestions on an earlier draft of this paper.*

John Esposito has taught EFL in Japan for six years. He is currently a doctoral student at the University of Colorado at Boulder.

References

Ballard, B., & Clanchy, J. (1991). Assessment by misconception: Cultural influences and intellectual traditions. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 19-35). Norwood, NJ: Ablex.

Beichman, J. (1986). *Masaoka Shiki*. Tokyo: Kodansha International

Blyth, R.H. (1981). *Haiku: Eastern culture, Vol. 1*. Tokyo: Hokuseido Press.

Carrell, P.L. (1987). Text as interaction: Some implications of text analysis and reading research for ESL composition. In U. Connor & R.B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 47-55). Reading, MA: Addison-Wesley.

Carson, J.E., Carrell, P.L., Silberstein, S., Kroll, B., & Kuehn, P.A. (1990). Reading-writing relationships in first and second language. *TESOL Quarterly, 24*(2), 245-266.

Carson, J.G. (1992). Becoming biliterate: First language influences. *Journal of Second Language Writing, 1*(1), 37-60.

Clancy, P.M. (1986). The acquisition of communicative style in Japanese. In B.B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (pp. 213-249). New York: Cambridge University Press.

Coe, R.M. (1987). An apology for form: Or, who took the form out of the process? *College English, 49*(1), 13-28.

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing*. Cambridge: Cambridge University Press.

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning, 39*(1), 81-141.

D'Angelo, F. (1980). *Process and thought in composition*. Cambridge, MA: Winthrop.

Friedlander, A. (1990). Composing in English: Effects of a first language on writing in English as a second language. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 109-125). New York: Cambridge University Press.

Grabe, W., & Kaplan, R.B. (1989). Writing in a second language: Contrastive rhetoric. In D.M. Johnson & D.H. Roen (Eds.), *Richness in writing: Empowering ESL students* (pp. 263-283). New York: Longman.

Hare, R.M. (1982). *Plato*. New York: Oxford University Press.

Harris, D.P. (1990). The use of "organizing sentences" in the structure of paragraphs in science textbooks. In U. Connor & A.M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives* (pp. 67-86). Alexandria, VA: TESOL.

Henderson, H.G. (1967). *Haiku in English*. Tokyo: Charles E. Tuttle.

Higginson, W.J. (1985). *The haiku handbook: How to write, share, and teach haiku*. Tokyo: Kodansha International.

Hinds, J. (1982, May). *Contrastive rhetoric: Japanese and English*. Paper presented at the Annual Convention of Teachers of English to Speakers of Other Languages, Honolulu, HI.

Hinds, J. (1987). Reader versus writer responsibility: A new typology. In U.

Connor & R.B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 141-152). Reading, MA: Addison-Wesley.

Hinds, J. (1990). Inductive, deductive, quasi-inductive: Expository writing in Japanese, Korean, Chinese, and Thai. In U. Connor & A.M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives* (pp. 87-109). Alexandria, VA: TESOL.

Hinkel, E. (1994). Native and nonnative speakers' pragmatic interpretations of English texts. *TESOL Quarterly, 28*(2), 353-376.

Hirose, K., & Sasaki, M. (1994). Explanatory variables for Japanese students' expository writing in English: An exploratory study. *Journal of Second Language Writing, 3*(3), 203-229.

Horiuchi, T. (1993, July 8). When a poem's form and content kiss, it's bliss. *The Japan Times,* 15.

Horowitz, D.M. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly, 20*(3), 445-462.

Jenkins, S., & Hinds, J. (1987). Business letter writing: English, French, and Japanese. *TESOL Quarterly, 21*(2), 327-345.

Kaplan, R.B. (1966). Culural thought patterns in intercultural education. *Language Learning, 16*(1), 1-20.

Kaplan, R.B. (1987). Cultural thought patterns revisited. In U. Connor & R.B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 9-20). Reading, MA: Addison-Wesley.

Kaplan, R.B. (1988). Contrastive rhetoric and second language learning: Notes toward a theory of contrastive rhetoric. In A.C. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric* (pp. 275-304). Newbury Park, CA: Sage.

Kaplan, R.B. (1990). Writing in a multilingual/multicultural context: What's contrastive about contrastive rhetoric? *The Writing Instructor, 10*(1), 7-18.

Knoblauch, C.H., & Brannon, L. (1984). *Rhetorical traditions and the teaching of writing.* NJ: Boynton/Cook.

Kobayashi, H. (1984). Rhetorical patterns in English and Japanese. *TESOL Quarterly, 18*(4), 737-738.

Krapels, A.R. (1990). An overview of second language writing process research. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 37-55). New York: Cambridge University Press.

Lay, N. (1982). Composing processes of adult ESL learners: A case study. *TESOL Quarterly, 16*(3), 406.

Lefkowitz, N. (1987). *From process to product: Beginning-intermediate writing skills for students of ESL.* Englewood Cliffs, NJ: Prentice Hall Regents.

Leki, I. (1991). Twenty-five years of contrastive rhetoric: Text analysis and writing pedagogies. *TESOL Quarterly, 25*(1), 123-143.

Leki, I. (1992). Understanding ESL writers: a guide for teachers. Portsmouth, NH: Boynton/Cook.

Liebman, J.D. (1992). Toward a new contrastive rhetoric: Differences between Arabic and Japanese rhetorical instruction. *Journal of Second Language Writ-*

ing, 1(2), 141-165.

Liebman-Kleine, J. (1986, March). *Toward a contrastive new rhetoric–A rhetoric of process*. Paper presented at the Annual Convention of the Teachers of English to Speakers of Other Languages, Anaheim, CA.

Lindemann, E. (1982). *A rhetoric for writing teachers*. New York: Oxford University Press.

Martin, J.E. (1991). Contrastive rhetoric: Implications of a revised approach to text. (ERIC Document Reproduction Service No. ED 329 118)

Matalene, C. (1985). Contrastive rhetoric: An American writing teacher in China. *College English, 47*(8), 789-808.

Mok, W.E. (1993). Contrastive rhetoric and the Japanese writer of EFL. *JALT Journal, 15*(2), 151-161.

Ooka, M. (1991). *The colors of poetry: Essays on classic Japanese verse*. (T.U. Lento & T.V. Lento, Trans.). Rochester, MN: Katydid Books.

Purves, A.C. (1988). Introduction. In A.C. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric* (pp. 9-21). Newbury Park, CA: Sage.

Purves, A. & Hawisher, G. (1990). Writers, judges, and text models. In R. Beach & S. Hynds (Eds.), *Developing discourse practices in adolescence and adulthood* (pp. 183-199). Norwood, NJ: Ablex.

Raimes, A. (1991). Out of the woods: Emerging traditions in the teaching of writing. *TESOL Quarterly, 25*(3), 407-430.

Roberts, W. R. & Bywater, I. (Trans). (1984). *The rhetoric and the poetics of Aristotle*. New York: Random House.

Rodgers, P.C. (1965). Alexander Bain and the rise of the organic paragraph. *Quarterly Journal of Speech, 51*, 399-408.

Sato, H. (1983). *One hundred frogs: From* renga *to haiku to English*. New York: Weatherhill.

Saunders, A.N. (1970). *Greek political oratory*. London: Penguin Books.

Scarcella, R.C. (1984). How writers orient their readers in expository essays: A comparative study of native and non-native English writers. *TESOL Quarterly, 18*(4), 671-688.

Severino, C. (1993). The "doodles" in context: Qualifying claims about contrastive rhetoric. *The Writing Center Journal, 14*(1), 44-62.

Shearer, N.A. (1972). Alexander Bain and the genesis of paragraph theory. *Quarterly Journal of Speech, 58*, 408-417.

Silva, T. (1990). Second language composition instruction: Developments, issues, and directions in ESL. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 11-23). New York: Cambridge University Press.

Strevens, P. (1987). Cultural barriers to language learning. In L.E. Smith (Ed.), *Discourse across cultures: Strategies in world Englishes* (pp. 169-178). Englewood Cliffs, NJ: Prentice Hall.

Tsujimura, A. (1987). Some characteristics of the Japanese way of communication. In D.L. Kincaid (Ed.), *Communication theory: Eastern and Western per-*

spectives (pp. 115-126). San Diego, CA: Academic Press.
Ueda, M. (1967). *Literary and art theories in Japan*. Ann Arbor, MI: University of Michigan, Michigan Center for Japanese Studies.
Yamada, K. (1993). Visions: An academic writing text. *JALT Journal, 15*(1), 115-118.
Zamel, V. (1982). Writing: The process of discovering meaning. *TESOL Quarterly, 16*(2), 195-209.

(Received April 14, 1996; revised January 23, 1997)

Appendix: Haiku and Explanatory Paragraphs Written by the Author's Japanese EFL Students

A beautiful sandy beach
A little turtle under an evening glow
Moves alone to a new world

It is the sight of a beautiful sandy beach which I have never seen before. A little turtle who came into existence just now tries to walk to an unknown world (ocean) under an evening glow. From now on, I think that I will try to come to the unknown and new world with my hope by myself.

(T. E.)

In the winter woods
Many trees are standing silent
My heart feels loneliness

I was walking in the winter woods. The cold wind was blowing. And it was snowing. The winter woods are very silent. No noise. I noticed that many animals are hibernating. And I noticed that there are only many trees. Many trees are standing silent. The silence made me think of loneliness.

(T. M.)

The clouds in the clear sky
Float in the wind
Unsettled like people's minds

When I was walking cold outside, on the spur of the moment I looked at the sky. It was so clear, blue, and beautiful. There are many white clouds. The clouds made me think of my memories, promises with somebody a long time ago because we all share this sky. And I noticed that the state of clouds is not everlasting. The clouds made me think of people's unsettled minds.

(M. Y.)

Reviews

Teachers as Course Developers. Kathleen Graves (Ed.). Cambridge: Cambridge University Press, 1996. 213 pp.

Reviewed by
Greta J. Gorsuch
Mejiro University

Anyone who attended presentations at JALT '94 in Matsuyama or heard the JALT '94 plenary address, "Teaching as Decision Making: A Means to Reflective Practice," by Don Freeman, TESOL President at that time, will recognize the general theme of Graves' *Teachers as Course Developers*. While Freeman called for recognition of a theory of teaching, Graves asks teachers to create theories about their own teaching by focusing on their own experiences. In Graves' book, this theory building takes place against the backdrop of individual teachers developing courses.

The book contains eight chapters, the first two of which Graves wrote as introductions to the book and to her framework of course development. The other six are accounts of course development written by ESL and EFL teachers practicing in various situations—a man in Boston helping Chinese immigrants learn essential English as workers in a nursing home, a woman in Japan designing a social studies course for seventh graders in an international school, a woman in Ecuador creating an academic English course for adults, a woman helping students become better writers of English in Brazil, another woman in Japan designing an advanced listening course for Japanese junior college students, and a woman in the U.S. helping Asian, European, and Latin American executives use English more effectively in the corporate world. Reading these accounts just for the feel of the variety of situations in which ESL/EFL teachers teach is reason enough to get the book. The experiences of these teachers are simply fascinating.

At the end of each teacher's account, Graves adds an "Analysis and Tasks" section which help individual readers/teachers focus on various issues raised by the teacher's account. In plain, accessible language, Graves calls on teachers, for example, to create an explanation of goals and objectives for a course they teach to non-teachers, and then to

JALT Journal, Vol. 19, No. 2, November, 1997

note how they do the explaining and what elements stand out in their explanations. Anyone who has tried this with department heads at his or her own school will know how very painful, yet clarifying and revealing this process can be. An added bonus is the inclusion in apppendicies of many of the documents the contributors created in the process of developing their courses. In particular, Carmen Blyth (chapter five, pp. 86-118), provides a detailed daily syllabus of her English for Academic Purposes course which should interest other teacher-cum-course developers.

Graves' framework of course development (described in chapter two) consists of needs assessment, determining goals and objectives, conceptualizing content, selecting and developing materials and activities, organizing the materials and activities, evaluation, and consideration of resources and constraints. Each of the teachers' accounts in chapters three through eight focus on some aspect of these framework elements. Johan Uvin (chapter three, pp. 39-62), for instance, focuses on needs assessment, an issue in curriculum and course development that can hardly be explained or described enough. The elements of Graves' framework of course development are classic (see, for example, Brown, 1995, p. 20), but she goes into more detail than do many curricularists, which makes *Teachers as Course Developers* all the more accessible. Of particular interest is Graves' "conceptualizing content," where she helps teachers clarify their assumptions about language, language use, syllabuses, classroom activities, and learning strategies through a kind of expanding visual grid.

Along with aids such as this grid, through diary writing, and through a process Graves calls "problematizing," which really is a kind of hypothesis making, teachers can develop courses and thus, in my mind, create theories of their own teaching. However, Graves never states explicitly that this is what she is doing. Instead, she uses terms like "draw on their own experience," "providing them with a conceptual framework," "identify challenges," "figure out," and "need to understand," to describe what she thinks teachers should be doing in their journeys inwards.

Graves' approach to theory lacks conceptual clarity in that she makes a distinction between what she calls "theory in the general sense," and "personal theory." She seems to believe that the theory in her "theory in the general sense" is not a product of human cognitive processes and human subjectivity, and that "personal theory" is. She cites Prahbu (p. 2) in defining theory in the general sense as "an abstraction that attempts to unite diverse and complex phenomena into a single principle or system of principles." She then defines personal theory as "a subjective understanding of one's practice . . . that provides coherence and direction" (p. 2). I fail to see the difference between the two definitions. Isn't "one's practice" a set of "diverse and complex phenomena"?

Based on this dubious distinction between theory and personal theory, Graves seems to say that theories coming from sources external to the teacher are to be ignored, while theories that teachers themselves create in the course of teaching are to be the sole focus. One example of this comes from Graves' account of a teacher who was assigned to teach a 140 student conversation class. In the context of discussions of doing a needs analysis questionnaire with these students, the teacher commented she wanted to see some examples of needs analysis instruments that others had done so she herself could get an idea of what kind of needs analysis she wanted to do. Graves took exception to this because she felt the teacher had to "problematize" her situation first, that is, the teacher had to "understand the givens of her situation . . . identify the challenges that will shape her decisions . . . and figure out what must and can be done" (p. 5). She seems to say that the teacher has to first create some of her own hypotheses and theories about her situation before consulting external sources, such as books, articles, or colleagues. But who is to say that perusing needs analysis instruments written by others is not part of this particular teacher's theory building? Perhaps by seeing what others had done (theorized) in their situations, the teacher could more effectively conceptualize the whole notion of doing a needs analysis. Taken several steps further, why should teachers be reading *Teachers as Course Developers*? The theories developed by the contributors to the book are, after all, external to the readers of the book. Why should their stories matter to me, for instance, if I am to build my own theories about my course development processes, without external influences? No doubt this is stating it too strongly, but it does illustrate the puzzling loop I perceive in Graves' thinking.

Despite this lack of clarity about the uses of theory, Graves has created an impressive volume of teachers' stories, and has helped them document the processes of their teaching. Perhaps by reflecting on their stories, we can more effectively understand our own.

Reference

Brown, J.D. (1995). *Elements of language curriculum*. Boston: Heinle & Heinle.

Verbal Hygiene. Deborah Cameron. London: Routledge, 1995. 264 pp.

Reviewed by
Virginia LoCastro
International Christian University

Readers of the *JALT Journal* will know Deborah Cameron as the author of *Feminism and Linguistic Theory* (FLT) (1985), in which she addressed the theory of the relation between language and world view as seen in the relation of language and gender. Whether or not one agrees with it, FLT cannot be dismissed, as such critical and thought-provoking syntheses are unfortunately rare and are thus to be welcomed, irrespective of the reader's ideology.

It was with these thoughts in mind that I greeted her 1995 work, *Verbal Hygiene* (VH) . Cameron brings her familiar intellectual honesty and passion to bear on the subject of "verbal hygiene," which she defines as "the practice by which people attempt to regulate" language use (p. XX). She coined this expression to cover a variety of evaluative activities people engage in to combat what they view as abuses of "our" language. One's first reaction to the phrase "verbal hygiene" may be to conjure up images of an individual, perhaps oneself, sitting in a language laboratory or in front of a mirror, carefully repeating after tapes in order to get one's French "r" out right–pronounce it "cleanly." Or more locally, an image of new female employees at Japanese companies who undergo special training so that they can start on April 1st, able to use *keigo* appropriately as well as serve tea, answer the phone, and dress, and even sit down "politely." Clearly, this 264 page book is about more than such matters; yet such personal images provide anecdotal evidence of the everyday discourses and practices which fit into what Cameron means by verbal hygiene.

VH focuses on language use, rather than usage. "Usage" refers to the conventionalized, generally accepted "rules" about correctness, which are typically found in dictionaries and pedagogical grammars of a language. If a student asks me whether "If I were you . . ." or "If I was you . . ." is correct, for example, I always explain that for an English examination only the first one is correct usage. The second one, "If I was you . . ." involves the question of use, that is, what people actually do in everyday situations. It is this area of linguistic analysis that concerns Cameron. The importance of use is apparent in the ubiquitous presence of writers devoting whole books to the state of a language, usually their mother tongue, a well-known example of such activity, for Americans, being William Safire's syndicated column "On Language."

However, VH tackles more than instances of language use or misuse. First, it makes a spirited survey of contemporary verbal hygiene practices, such as the political correctness movement in the U.S. Second, drawing from current language and cultural theory, it analyses the motives and meanings underlying verbal hygiene. Third, VH addresses linguists directly, many of whom take pride in what they believe to be the description of language use, reject any notion of prescriptivism, and hold firm to the stance that change in language use and usage is normal and inevitable. Although a linguist herself, Cameron argues otherwise.

In chapter 1, "On verbal hygiene," Cameron argues that investigation of the phenomenon of verbal hygiene is a worthy pursuit for its ability to shed light on the relation between language, society, and identity. Starting with the observation that "humans do not just use language, they comment on the language they use" (p. 1), she contends that normativity is an essential part of language-using, which is a "social, public act." This implies the need for minimum normative standards, in order for communication with a minimum of problems to occur between and among individuals. Distinguishing verbal hygiene, norm-observing, from prescriptivism, which seeks to enforce norms, the first chapter addresses the social construction of normativity and the underlying ideology of value judgments. Cameron considers the ways the fear of fragmentation of communication covers deeper fears of social fragmentation deeply embedded in post-modern societies (Turner, 1989). Having established that verbal hygiene is essentially about values, Cameron next examines particular sets of practices and values underlying evaluative discourse about language for writers (chapter 2), national educational curricula (chapter 3), political correctness (chapter 4), and gender and language (chapter 5). In all cases, her concern is not about using "proper" grammar for its own sake; "proper" language use has symbolic meanings at the individual and societal levels.

"Restrictive practices: The politics of style," chapter 2, examines institutionalized verbal hygiene practices of style guides for writers, specifically journalists, focusing mostly on *The Times* (of London) and *The Chicago Manual of Style*. This chapter shows the moral judgments underlying the myth that "good" writing is self-evident. A careful study of the style manuals along with interviews of editors in the UK and the US enables Cameron to describe and explain the role of craft professionals, particularly editors, in regulating the language of the print media. This leads her to conclude that the entire endeavor is "characterized by authoritarianism, mystification, irrationality, and lack of critical engagement" (p. 77). Editorial fetishes of "correctness," "consistency," "transparency," and "uniformity" communicate a preoccupation with the perfectibility of communication and hyper-standardization. Cameron shows, moreover, that the underlying purpose is to comodify "style," to sell it as a high-class product, and one which gate-keepers such as editors control. While criticizing the verbal hygiene practices of the style keepers, Cameron is more interested in raising awareness and demystifing the workings of this particular form of VH.

Chapter 3, "Dr Syntax and Mrs Grundy: The great grammar crusade," is a case study of the "curriculum wars" which led to the 1988 Education Reform Act and the national school curriculum in the UK. It will be immediately accessible to British readers, American readers will relate it to debates in the U.S. about a "national curriculum." Readers based in Japan will note the Ministry of Education's (MOE) continuing involvement in language education at all levels, as evident in recent news reports on the introduction of English in primary schools. Among the controversies subsumed within the British debate about a national school curriculum was a highly politicized one about English teaching, specifically the teaching of grammar. The "pro-grammar ideologues" (p. 86) held classroom teachers responsible for "falling standards among pupils and ideological subversion among teachers" (p. 89). Thus, under the cover of emphasis on proper spelling and grammar, the conservative supporters of the National Curriculum proposal insisted on standard English as the only acceptable dialect, for they feared threats to the mother tongue and national culture. Cameron claims that grammar became a "moral metaphor" for a "cluster of related political and moral terms: order, tradition, authority, hierarchy, and rules" (p. 95).

As with editors' verbal hygiene, the moral symbolism becomes apparent upon examination: verbal hygiene and moral or social hygiene cannot be separated and, while one can deplore the beliefs which inform the value judgments, one cannot ignore the apparent importance of the standards and values being promoted or the power of those holding them.

Chapter 4: "Civility and its discontents: Language and 'political correctness'" approaches a highly charged topic. As Cameron states, the political correctness (PC) debate is essentially about deciding whose values should be conveyed through planned, pro-active efforts to change language use. On one level, opposition to politically motivated language change represents rejection of feminism, multiculturalism, and other minority group issues; yet, on a deeper level, it signals questions about the extent to which language can influence ideas and about folk linguistic views of how meanings are created and by whom. This chapter briefly surveys the development of PC, its origins in the New Left in the U.S. in the 1960s and 1970s, and the non-sexist language guidelines developed by the University of Strathclyde's Programme of Opportunities for Women Committee (POWC). Cameron explains that the arguments in favor of using non-sexist language include the notion of "civility" (PC language purports to show greater sensitivity to others' feelings); a concern with accuracy and transparency of meaning (generic masculine terms, for example, might be misleading); and finally fairness (inclusiveness in language use of men and women). Clearly, the verbal hygiene practices of PC can have a long-term effect on changing attitudes. Average people, who may or may not be interested in social change, often object to this. They zero in on what they see as an organized attempt to destroy the existing relationships between words and reality, following the commonsense notion that words correspond one-to-one to things in the "real world." Specifically, Cameron states "the debate on 'political correctness' is most obviously a debate about how democracies made up of diverse populations subscribing to a variety of beliefs and customs are to preserve a common culture" (p. 160). The endless discussion about language implies a lack of social consensus and the end of the belief in a value-free language. Consequently, Cameron advocates public acknowledgment about how language is used and who decides how it is to be used.

The final topic-oriented chapter, chapter 5: "The new Pygmalion: Verbal hygiene for women," discusses the self-improvement movement's concern with a linguistic remodeling of the individual. Verbal hygiene in this case focuses on the notion that the way women speak is problematic, particularly in male/female communication in the work place where men view women as lacking appropriate management skills. A rich body of literature exists on language and gender, in which linguistic research comes together in the best-sellers of Tannen (1986, 1990).

One consequence of this concern has been to advise women on how they should speak and perform their identities as ideal women in the workplace. Cameron illustrates this phenomenon by drawing on both historical and contemporary advice literature. She comments that the proliferation of this in recent years may particularly reflect modern insecurities about femininity. Further, citing an article in *Cosmopolitan*, she notes the recognition by the general public that female speech habits may not be helpful and that learning to be more assertive might enable a women to function more effectively. Assertiveness training subsequently became the main thrust of the self-help literature and workshops have become part of the mainstream efforts to empower women, linguistically and otherwise.

However, the story does not end here as Cameron is quick to point out. Career advice for women that they should talk like men (e.g. speak more directly, avoid using tags, hedges, and interrogative intonation on declaratives) conflicts with relationship advice offered in many of the same magazines, where "feminine" interpersonal skills are valued. Efforts to overcome the negative stereotype of career women as lacking in authority and credibility lead women in exactly the opposite direction of the socially approved norms for women who wish to be attractive to men. Once again verbal hygiene practices go beyond a manipulation of linguistic features; in this case, women seek to overcome their insecurities both in the workplace and in relationships by resorting to the advice in the self-help manuals, whose main function is maintenance of male-female distinctions.

The final chapter, chapter 6, "On the state of the state of language," synthesizes the issues raised in the previous five chapters. The chapter's title refers to attempts to clarify the nature of language, an "overview that will tell us where we are" (p. 212), as if language were something static and unchanging. In VH, however, Cameron demonstrates how "the 'state of language,' is a discursive construct, not an objective description of certain linguistic phenomena" (p. 212). She contrasts the ubiquitous view of language as a "natural" phenomenon with language-using as social practice, as a form of behavior through which human beings act in the world and suggests some principles informing the concerns of verbal hygienists.

Finally, she returns to the question of the role of experts, specifically how linguists can meet the challenge of verbal hygiene without compromising their intellectual values. As Cameron, in my opinion, rightfully states, linguists "make value judgments about language use without stating their criteria" or acknowledging that they are doing so. If linguists claim a particular instance of language use is "acceptable" or "appropriate," they need to define their terms and not hide behind a false cover of "anything goes," as if any use of language were equally acceptable in the social world; normative intent underlies any statement of appropriacy. Cameron concludes by calling on linguists to acknowledge the deep-seated concerns of those who support verbal hygiene, and to work with them, rather than denying what seems to be a pervasive, human phenomenon.

VH has obvious relevance for readers with interests in sociolinguistics, the sociology of language, critical discourse analysis, and the study of language, politics, and ideology. While extremely well-written and edited, and full of interesting, compelling examples, it is not an easy read as, to make her points, Cameron's argumentation on specific points can continue through an entire chapter, while the macro level argument, i.e. that verbal hygiene is about a lot more than promoting "correct" language use, is sustained throughout the entire book. There are numerous treasures along the way: humorous anecdotes, well-chosen phrases, curious examples of verbal hygiene, and indeed the carefully worked out argumentation. One has to admire the author for her intellectual honesty in taking apart the verbal hygienists as well as her own colleagues, fellow linguists.

As for the usefulness of VH for *JALT Journal* readers, I highly recommend the book to readers who want a better understanding of what human beings do with language in the real world. VH provides interesting and timely reading whether one wants an engaging review of current verbal hygiene practices or an exercise in cultural analysis of one fascinating manifestation of what are generally viewed as conservative reactions to language use in postmodern societies.

However, Cameron's conclusion that linguists should descend into the fray and address the perceived needs of lay people to have some control over what happens to language in society most directly relates to the current situation in Japan regarding the possible introduction into the primary schools of English language education. Recent deliberations inside the MOE concerning the wisdom of introducing English language education into the primary schools in Japan are clearly not an example of verbal hygiene. However, the same critical analysis which Cameron employs indicates that teaching English as a foreign language at an earlier age symbolizes an attempt to do more than have the pupils learn some English. It is an experiment in social engineering, with one of its goals to help pupils develop their ability to express themselves in any language, including their own language, as some of the statements of Japanese people imply in private discussions of this matter. Articles in the popular press already indicate the controversy greeting the MOE report on this topic (see Fukushima, 1996). It is early and these articles do not provide enough detail to get a clear picture of the MOE's stance; nevertheless, it is clear that popular opinion is based on folk linguistic beliefs which, from a linguistic point of view, may not be strictly valid. Linguists and other experts in education have a clear role to help clarify the extent to which many of the issues regarding English language education in primary schools concern language and education less than they deal with social, political, and economic issues, and deeply involve the ethnolinguistic identity of Japanese people. Linguists need to work with the lay public to develop more informed standards, to make informed decisions. Perhaps Cameron's book can help us make sense of this particular effort to regulate language and society.

Acknowledgment

I'd like to thank Robin Antepara for her helpful comments on this book review.

References

Cameron, D. (1985). *Feminism and linguistic theory*. Basingstoke, UK: Macmillan.

Fukushima, K. (1996, May 13). Our word for today is 'controversy.' *Asahi Evening News*, 7.

Tannen, D. (1986). *That's not what I meant! How conversational style makes or breaks your relations with others*. New York: William Morrow.

Tannen, D. (1990). *You just don't understand: Women and men in conversation*. New York: Ballantine.

Turner, B.S. (Ed.) (1989). *Theories of modernity and postmodernity*. London: Sage Publications.

Understanding Metaphor in Literature:An Empirical Approach. Gerard Steen. Longman. London. 1994. xiii + 263 pp.

Reviewed by
Valerie Fox
Sophia University

This dense book, part of Longman's *Studies in Language and Linguistics* series, attempts to provide empirical evidence to support several hypotheses about how people understand metaphor in literature. Steen has been greatly influenced by the "cognitive turn" in metaphor analysis; his inquiries are meant to "make progress from the recently achieved theoretical perspective on metaphor as cognition to the development of a cognitive view of metaphor in discourse processing" (p. 5). In the first chapter he reviews this achievement, in order to provide a firm foundation on which to build the reporting of his research. In chapter one he also lists the questions that constitute the main subject of the study:

If people's use of metaphor has become part and parcel of our view of cognition, and its proverbial relation to literature has been undermined, what is the relation between metaphor and literature? Can we still speak of such a thing as "literary metaphor"? And do metaphors in literature have a special cognitive function which can be differentiated from the cognitive function of metaphors elsewhere? Where do we have to look to find an answer to these questions: in language, in cognition, or still other areas related to literariness? (p. 5)

Steen's book contains three sections, each combining a review of pertinent articles with discussion of his own research. Each section owes much to a specific researcher or team. Part I, "Reader, Text, Context," builds on yet disagrees with some 'findings and methods found in Lakoff and Johnson's (1980) *Metaphors We Live By*. Part II, "Processes," builds on the work of Van Dijk and Kintsch (1983). Steen characterizes metaphor processes more specifically than these writers, paying more attention to communicative processes. Part III, "Properties," draws on and attempts to improve on Schmidt's empirical study of literature (1980).

Steen utilizes a variety of materials and procedures. Those less familiar with psycholinguistics and the empirical study of literature, but familiar with the analytical style of Lakoff, and Lakoff's work with Johnson, may find the contrast provocative. Lakoff and Johnson (1980) argue that metaphors in everyday language and life give coherence to life, and "play a central role in the construction of social and political reality" (p. 159). Their conclusions are analytical and speculative, based largely on analysis of lists of metaphorical usage. For example, they list and examine usage that are built upon ideas such as "ideas are food" or "the eyes are containers of the emotions" (pp. 46-50). Steen has gathered empirical evidence, thus taking into account factors that might be neglected by this more analytical approach.

Steen's array of techniques in and of itself will show some readers the vast potential for research in the study of reading. Techniques includes subjects' identification and explanation of metaphors in 400-word reading passages; rating of metaphorical phrases (highlighted within sentences and brief passages) using Semantic Differential pairings such as original-trite, shocking-touching; and, comparison of subjects' performance during "thinking aloud" experiments. I will describe, generally, most of Steen's experiments, as the variety of methods is a strength of this book.

In the first of two experiments contained in his chapter "Metaphor and Literariness," Steen asked his subjects to underline and explain ten metaphors in a literary excerpt: Norman Mailer's *Miami and the Siege of Chicago: An Informal History of the American Political Conventions of 1968*. (This book blends a literary and subjective style with journalistic reportage.) In their explanations, 42 subjects were supposed to explain why the underlined passages were, in their opinions, literary or literature. There were two groups of subjects. The group which had recently received formal training in literary analysis underlined more metaphors than did those who had not recently received instruction. Steen also found that both groups tended to underline and explicitly identify metaphors that had been defined (by language and literature teachers) as having a "high degree of metaphoricity" as opposed to metaphors that had been defined as having a "low degree of metaphoricity." Explicit identification was determined by the subjects' use of words such as "metaphor," "image," and "analogy" in their explanations (pp. 59-61).

In a related experiment, Steen looks at some other effects of context on reading. He asked subjects to read two passages and underline and explain language that was literary or journalistic. Both excerpts were in Dutch, one having been taken from a daily newspaper, and the other one from a novel. The subjects were divided into two groups. One group was given the genuine identification of one passage as journalistic and the other as literary; the second group was told that the journalistic excerpt was literary, and the literary excerpt was journalistic. Thus, Steen could study the effects of "text presentation" (p. 67). He presents statistics to support his hypothesis that it is "the literary reading task which promotes attention to metaphors" (p. 70). Subjects highlighted metaphors as typically literary when they thought they were reading literature. They did not use expressions related to metaphors (metaphor, image, comparison) when they explained why language was typically journalistic.

In the book's second section, "Processes," Steen builds on the previously mentioned experiments. He incorporates aspects of other studies of reading, especially Van Dijk and Kintsch (1983), whose work he admires but considers limited as it focuses on the goal-directedness of reading. Steen will take into account "the role of social contexts in the formation of goals for discourse processing" (p. 85). His aim is to "develop an integral theoretical view of metaphor processing during reading" (p. 83). He then will use empirical research to investigate his "provisional picture" of metaphor processing (p. 83). He wishes to provide tools, as it were, that will facilitate the study of metaphor. To this end, he modifies Kintsch's model (Steen, p. 85) which he believes relies too much on the structure of the text as the basis for a theoretical model. The improved model should include three steps: decoding, conceptualization, and communication (p. 85). "Communication" aspects of cognitive processing have been particularly neglected, according to Steen.

In the literature review that precedes his experiments on process, Steen asks and provides possible answers for these questions: "Are all metaphors understood in two stages?" and, "Are metaphors always recognized as such?" (pp. 90, 94). Steen agrees with Gerrig that in time-limited situations (conversation, theater), there is minimal connection between comprehension and appreciation (p. 105). Bearing in mind the three-part model of reading, the many-faceted process of understanding metaphor, and the relevance of time limitations, Steen proceeds to use a thinking-aloud experiment to examine closely the process of metaphor comprehension. In so doing, he tries to determine whether the previously explained parts of the metaphor understanding process are important in literary reading, and also, whether there might not be additional factors that must be included in studies of literary metaphor processing (p. 107).

In the study, sixteen subjects (seven students and nine lecturers) were asked to read one of eight possible Dutch texts (seven literary and one non-literary). They were given the texts one sentence at a time. Each new, underlined sentence was added to the passage, on a new piece of paper. Readers were asked to verbalize everything that came into their minds. They were asked to concentrate on each new sentence, and not to re-read, although it was possible to do so. Also, subjects were told to refrain from explanation and interpretation.

Steen notes various kinds of processing that occur in the readers' responses. These include focus processing, vehicle construction, analogizing, functionalization, and refunctionalization (pp. 124-128; these terms are all clearly explained). Thus, Steen provides evidence to support a complex model for metaphor processing.

He concludes that some types of processing seem more likely to occur when reading literary texts, specifically those types of processing that he has defined as not involving analysis and explication. He suggests that these kinds of analysis are "probably much more tied to educational or scientific analysis than to other kinds of text processing" (p. 130).

Next, having classified aspects of metaphor processing, Steen uses some materials from a previous experiment (the two Dutch texts from his underlining experiment) to test his hypothesis that literary socialization influences metaphor processing. He compares the behavior of anthropologists and literature lecturers, and offers explanations for similarities and differences (pp. 151-154).

In the third section of this volume, "Properties," Steen builds on his findings regarding the nature of metaphor processing. In this section he classifies literary and journalistic metaphors according to five dimensions. Three are relatively cognitive: linguistic form, conceptual content, and communicative function; two classifications are "less cognitive": emotive value and moral position (p. 181). Steen opposes some literary critics' "devaluation of the text as an autonomous phenomenon" and their sometimes "extremely reader-oriented view of reading," (p. 165). The five dimensions reflect this broader view.

In two studies Steen asked subjects to rate metaphors using the Semantic Differential technique. The goal was to use the ratings to help classify specific metaphors according to the five property classifications. Subjects rated words or phrases contained within short passages from literary and journalistic sources. The SD technique, originally developed to study vocabulary, involves the rating of metaphors using opposites; the scales were chosen to pertain to the five property classifications. "Shocking-touching," for example, pertains to the emotive property of metaphor. Subjects rated a metaphor on a scale of 1 to 5, depending on whether they found it shocking or touching. (A "3" was chosen if both or neither was thought to be appropriate.) This technique has the obvious advantage of producing numbers that can be incorporated into easily understood tables and graphs.

One conclusion Steen draws from this study is that journalistic metaphors are biased, possibly because this discourse's concern for "societal" interests. Literary metaphors "express a factually more disinterested, aesthetic attitude" (207). Steen himself mentions that a weakness of this study is that the reader might rate the entire passage, or context, instead of just the highlighted metaphor. It would have been easy, assuming Steen's examples of passages are representative, for readers to determine whether a passage was excerpted from a newspaper or a work of fiction. If the raters were rating the entire passage, Steen's conclusions about the properties of the metaphorical language itself are weakened. Steen's mention of this possible weakness demonstrates a noteworthy aspect of this volume that the author tries to qualify his findings when necessary.

Researchers from other disciplines, including critical discourse analysis, may be inspired to imitate Steen's methods. Norman Fairclough has appealed to researchers to include close textual analysis in their work (p. 208). Fairclough suggests that researchers could strengthen their work with a "three-dimensional view of discourse and discourse analysis," a view that includes "analysis of context, analysis of processes of text production and interpretation, analysis of text" (p. 211). Even a reader who prefers more general analysis will probably be impressed by the variety of Steen's methods, which take into account the first and third types of analysis suggested by Fairclough. Steen's attempts to refine and adapt methods might provide models to researchers trying to incorporate the methods of social science into their work.

Some readers might think that Steen's efforts to distinguish between literary and journalistic processing is narrow, as it does not seem to account for the "intertextual" qualities (to use Fairclough's terminology) that seem self-evident in many twentieth century texts. Steen hypothesizes that works of literature will exhibit qualities of "polyvalence" and "form orientation" and his studies bear this out (p. 35). But those who study journalistic writing, political speeches, advertisements, and other kinds of texts might argue that these too are sometimes polyvalent and form-oriented to a high degree.

The book title of *Understanding Metaphor in Literature* indicates that this work also will be of interest to literary theorists and critics. However, Steen's remarks regarding literary critics are often condescending. For example, when contrasting the literary critic with the "ordinary reader," he describes the critic's reading as unsuitable as a model for reading because it is influenced by his or her status as a paid professional who has unlimited time to analyze texts (p. 75). This condescends not just to the critic but also the so-called ordinary reader, whose readings might be similar to critics' readings under similar circumstances, that is given the time and resources. In another instance, Steen refers to the "language game of literary criticism" (p. 76). The language of such assessments, especially the use of words such as "game," implying an amusement, contrasts critics' work with the serious, objective endeavors of scientists, work obviously associated here with Steen's own procedures.

That said, those whose interest in literature is more aesthetic than empirical may find that many of Steen's conclusions seem plausible, and his findings may even corroborate their own experiences of processing metaphors. Steen's clearly explained research might well inspire further research. It may offer advice by way of example for researchers interested in how people understand metaphor.

References

Fairclough, N. (1995). *Critical discourse analysis*. London: Longman.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Schmidt, S.J. (1980). *Grundriss der empirischen Literaturwissenschaft, [vol 1]: Der gesellschaftlische Handlungsbereich Literatur*. Vieweg, Braunschweig/Wiesbaden.

Van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

The Functional Analysis of English: A Hallidayan Approach. **Thomas Bloor and Meriel Bloor. London: Arnold, 1995. 278 pp.**

Reviewed by
Wendy L. Bowcher
Tokyo Gakugei University

The study of grammar is seldom seen as a vehicle for studying social aspects of English. However, *The Functional Analysis of English: A Hallidayan Approach* provides the tools to do just that. This book demonstrates a way of looking at not just what can be done grammatically in English, but what is done and why. Bloor and Bloor claim that:

> Since a speaker's or writer's choice of words is constrained by the situation of utterance, and since words and groups of words take on special significance in particular contexts, the grammar must be able to account for the way in which the language is used in social situations. (p. 4)

This claim underpins functional grammar in the Hallidayan model.

The Functional Analysis of English offers an accessible introduction to Halliday's functional grammar. While this book focuses on developing a basic working knowledge and understanding of the grammatical system as proposed by Halliday, it is not just a grammar book. It provides readers with useful and succinct notes on some basic notions in Halliday's linguistic theory, historical influences on the theory, practical suggestions for English language teaching, and suggestions for further study and reading in the area. This makes the book particularly useful not only for students, but also for teachers who wish to see the insights a grammar can provide into the connection between language and the functions it serves in our lives.

The authors cover much of the content of Halliday's *An Introduction to Functional Grammar* (1994), albeit in a less complex manner, and occasionally direct the reader to relevant chapters of Halliday's book for further reading or for clarification of points. The first chapter introduces some basic notions in the theory, such as linguistic choice, the centrality of the analysis of authentic texts in developing linguistic description, and the notion of "rank." Chapter two explains functional labeling in a grammatical system as well as the place of class labeling. This provides a useful link between traditional grammatical terms such as adverb and adjective, and functional grammatical terms such as modifier or actor. The remaining chapters take the reader carefully through terminology, examples, and exercises in the grammar.

Halliday's book does not include exercises for the student, and so, Bloor and Bloor offer a practical, hands-on exploration, making it ideal an companion workbook to the Halliday volume. The exercises appear at the end of almost every chapter along with answers or suggested answers, depending on whether the question is open-ended or not. Exercises are varied, interesting, and challenging and make use of authentic texts in English from a variety of sources, including literature, the sciences, instruction manuals, recipes, jokes, and oral exchanges. They provide practice in analyzing clauses, identifying certain grammatical features of texts, comparing and contrasting texts, and interpreting grammatical choices in texts. The following exercise is taken from chapter six, "Process and Participant."

> Exercise 6.6
> Explain the following old joke in terms of Process and Participant.
> Comedian A: My dog's got no nose.
> Comedian B: Your dog's got no nose? How does he smell?
> Comedian A: Terrible. (p. 129-130)

The use of authentic texts throughout confirms the authors' aim to link grammar with the ways it is used in different situations, and offers a meaningful approach to the study of grammar rather than an approach which views grammar as merely a set of rules. The way in which the exercises and their answers are set out, however, is a shortcoming. The exercises appear on pages adjacent to the answers, making it difficult for students to work without referring to the answers. Because of this, the analyses in these exercises are probably best used as examples for analyzing other texts that the teacher or students bring to class. However, teachers do need to be aware that authentic texts may be more difficult to analyze than they first appear.

The second to last chapter of the book suggests possible applications of the grammar and outlines some significant research conducted within a functional framework. This chapter includes a section on English language teaching applications and a section on writing in science and technology. Both of these are minor sections in the book but they provide very useful insights. The section on English language teaching applications overviews some significant work within the Hallidayan (and related functional) framework on English language teaching, for example, cohesion, genre studies, and hedging in academic writing. It also provides examples of the application of a functional grammar in the TESOL classroom, specifically for the teaching of academic writing. One example refers to the way in which academic writers modify their claims. Bloor and Bloor note that:

> . . . when researchers writing in English make knowledge claims based on their research evidence, they rarely make bald confident statements,

> but they usually modify their propositions by the use of modal verbs such as 'may,' modal adjuncts such as possibly or lexical items that decrease the force of a proposition such as 'indicate' or 'appear.' (p. 231)

The section on writing in science and technology would be useful for anyone required to teach a content-based curriculum. In this section the authors not only direct the reader to relevant research in the field, but outline specific examples of the ways a functional grammar can lead to an understanding of the way in which language is used and structured in certain content areas. An example of this is the following quotation which refers to scientific writing. Bloor and Bloor state that:

> The tendency to use Nominal Groups rather than verbal processes has a number of major effects on scientific text. Firstly, it is a means whereby all reference to people can be omitted, and scientific knowledge can be presented as though it has some external objective reality quite apart from the people who are engaged in observing or researching it. This facilitates the expression of general "truths" and "claims" about the nature of the world. (p. 223)

The final chapter sketches the historical setting of Hallidayan approaches within the field of linguistics. This includes some of the major influences on Hallidayan linguistics and some of the differences between this approach and that of other linguists such as Chomsky. This is particularly useful as background information for language and linguistics teachers who are not particularly familiar with the Hallidayan school of linguistics and its connection to other schools of linguistics. It is also an accessible guide for students who are studying introductory linguistics.

Each chapter includes a summary of the main points covered and each chapter but the first includes a "further study" section directing readers to significant and related research in the area covered in that chapter. At the end of the book is a comprehensive glossary of terms, a full bibliography, and an index.

The authors are both highly experienced practitioners in the field of English language teaching, teacher training, and applied linguistics research. Within the pages of this book they demonstrate a thoughtful understanding of what teachers and students would like in a textbook on grammar, particularly if they are approaching this grammar for the first time.

The Functional Analysis of English: A Hallidayan Approach is suitable for introductory linguistics courses and English grammar courses at the university level, as a reference guide to functional grammar for language and linguistics teachers, as a companion workbook to Halliday's *An Introduction to Functional Grammar*, as a source book for functional grammar exercises using authentic texts, and for anyone with an interest in functional grammar.

Reference

Halliday, M.A.K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.

Teaching Business English. Mark Ellis and Christine Johnson. Hong Kong: Oxford University Press, 1994. xiv + 237 pp.

Reviewed by
Steffen Eckart
ALC Education, Kanagawa Branch

Imagining a more complete volume for the instructor motivated to improve his or her business English classes is a difficult task. Teaching Business English effectively synthesizes crucial information about professionals and their working world with theoretical principles drawn from teaching methodology, linguistics, and language testing. This mix of specific content issues with more general teaching concepts makes Ellis and Johnson's book ideal for those teachers of English for Business Purposes who are starting out, as well as experienced instructors who seek a fresh reference.

The book branches into three parts. Part one, "Introduction to Business English," draws upon the history of the field, the categories of learners who commonly need business English, the types of schools where English is taught, and resources available to the developing business English teacher. Part two, "Analyzing the needs of the learners," suggests means of gathering necessary information. Needs are broadly based on four learner characteristics: existing language abilities, job type such as managerial staff or technical staff, purposes for learning, and individual learner variables such as nationality and educational background. Included in this section are detailed charts and tables which break down these needs into categories. Chapter 9 is particularly useful because it connects business skills with language functions. Part three, "Activities and materials," guides the reader through available textbooks and offers detailed suggestions for creating original materials. The latter are grouped by Ellis and Johnson into two chapters. One covers framework materials, which are "diagrammatic representations which can be used to generate language" (p. 131), and the other examines authentic materials, which are treated as "any kind of material taken from the real world and not specifically created for the purpose of teaching" (p. 157). This third part provides concrete examples of how to implement the planning approaches which are detailed in the first two parts.

The book's push for teachers to carefully consider learner needs and to involve them directly in course design broadly stems from developments in Communicative Language Teaching (CLT) which have occurred over the past 25 years. In addition, advocates of CLT such as Savignon (1991, p. 266) have argued that program goals must be elaborated "in terms of functional competence." Examples in *Teaching Business English* include recommending and agreeing (p. 9). Mastering such functions depends upon a classroom where learners "feel secure and free of stress" (Ellis, 1994, p. 479), and where real communication is encouraged. This book is practical in nature, providing numerous ideas for implementing these principles, particularly in part three.

In a Business English course, the point of all this consideration of needs hinges on the teacher's intention to improve performance, defined by the authors as being "operationally effective" (p. 131). Language training must be carefully aimed in order to efficiently develop the language skills necessary for the target situations. Students will judge their teachers with the same expectations of professionalism that they would hold for other training programs conducted in their primary language. While it may be exciting for teachers to have students who are both highly motivated and who have specific learning objectives, it can also be a source of stress for the teacher who does not have the knowledge and materials needed to meet those objectives.

As a basic resource, *Teaching Business English* has only one particular omission that should be borne in mind. Little information is provided about Computer Aided Instruction and useful resources on the Internet for teachers and learners. Nevertheless, this book is a window of light onto a field with a lack of quality introductory books. I heartily recommend this volume to teachers and teacher trainers concerned with business.

References

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Savignon, S.J. (1991). Communicative language teaching: State of the art. *TESOL Quarterly, 25*(2), 261-277.

***Literacy and Script Reform in Occupation Japan: Reading Between the Lines*. J. Marshall Unger. New York: Oxford University Press, 1996. 176 pp.**

Reviewed by
David Cozy
Shonan International Women's Junior College.

Those attempting to teach *kanji* to Japanese children, or to foreigners studying Japanese, know how many hours of drudgery are necessary for their students to acquire competence. They must wonder, from time to time, whether it's worth it. Their students certainly do. Wouldn't it be easier, those struggling with *kanji* must feel, to simply use *romaji* or *kana* instead? By doing so, countless hours could be freed up for more worthy pursuits. These imaginary teachers and learners wouldn't be alone in their desire to simplify Japanese script. Calls to reform the Japanese writing system have been heard since at least the Meiji era and continue to ring out today. Literacy and Script Reform in Occupation Japan, which focuses on the immediate post-war years, is such a call.

Unger begins with a survey of script reform in the introduction and the initial chapters. (A more complete and less polemical overview can be found in Gottlieb, 1995.) In the course of this survey he encounters, and counters, many of the objections to script reform which have been advanced over the years.

The first, and perhaps the easiest for him to demolish, is the notion that *kanji* convey meanings, as opposed to speech sounds, or concepts directly, independent of how they are pronounced in the languages which employ them. *Kanji* express, according to this line of thought, concepts in a more immediate fashion than scripts such as the Roman letters you are now decoding, and are thus "unique among all forms of human writing" (p. 4). If this were in fact the case, one could understand why the Japanese would be unwilling to part with them. As Unger makes clear, however, this is not the case.

If each *kanji* really did express a unique idea or word, Unger argues, then "reading Chinese would be the same thing as . . . recalling the names of people while scanning a featureless list of telephone numbers; learning to read Chinese would be like memorizing the phone book for a town of several thousand customers" (p. 11). This, as he points out, is a feat few could manage. Millions of Chinese, though, do manage to read and write *kanji*, which suggests that kanji are not pure ideograms or logograms, and are thus not fundamentally different from scripts used elsewhere in the world.

That so many Chinese do learn to read, and that Japan has long been credited with one of the highest rates of literacy in the world, might seem to suggest that script reform, though in some ways desirable, is unnecessary. Unger believes, however, that the high rate of literacy assumed for pre-1945 Japan, and, indeed, for present-day Japan, is inflated (pp. 6-7).

One reason for this inflation is that literacy itself is a slippery concept. Does literacy in Japanese mean the ability to write and read one's name, or "productive facility in several socially prestigious and functionally distinct styles of Japanese and Sino-Japanese writing" (p. 25)? Most would probably say the dividing line between literacy and illiteracy should be placed somewhere between these two extremes, but exactly where is difficult to determine.

Unger believes that "in the early part of this century most Japanese possessed at best "a restricted set of skills that conferred only a portion of the liberating power we unthinkingly ascribe . . . to education" (p. 25). Indeed, even as late as 1948, a survey found that although complete illiteracy was negligible, only 6.2 percent of those participating were fully literate. A survey conducted in 1955-56 found that 50 to 60 percent of the participants lacked sufficient competence in written Japanese (p. 37).

Japan, therefore, was not as literate as some supposed it to be, and literacy was not evenly distributed: Men tended to be more literate than women, city people more literate than country people, retailers and artisan's more literate than fishermen and laborers (pp. 31-32). This might seem to suggest that lack of access to education, rather than the difficulty of Japanese script, lay at the heart of the problem, but Unger demurs. "Few Japanese," he concludes, "were totally illiterate, but the vast majority experienced some degree of difficulty in reading and writing that their education did not alleviate" (p. 43).

Unger continues his historical survey in the third chapter, "Script Reform from Within." As the title suggests, in this chapter the author argues that script reform is not, as some have claimed, a foreign notion imposed on Japan from outside. Rather, as Unger demonstrates through analysis of the historical record, script reform is something with which "thoughtful Japanese" (p. 44) have been concerned for centuries.

In chapter four, devoted to the role of SCAP (Supreme Command for the Allied Powers), Unger makes it clear that, even during the occupation when they could have, foreigners did not attempt to force the Japanese to reform their writing system (p. 59). They may, however, have facilitated such changes. The *toyo kanji* (*kanji* for daily use) list and other language reforms promulgated in 1946, while not imposed by the allies, were made possible by the eclipse of the rightists who had led Japan to defeat. By removing the conservatives from power, and promoting progressive politicians, SCAP opened a window of opportunity for reformers which had been more or less shut since the Taisho era.

This is ironic, because in detailing the behind the scenes machinations of SCAP bureaucrats, Unger demonstrates that several key players were lukewarm at best about script reform. One functionary who tried to alleviate his superiors' hostility to simplifying the Japanese language, the linguist Abraham Halpern, aware of ". . . the miasma of half-truths, speculation, [and] irrational and tangential reactions" (p. 128) that surrounded script reform, initiated an experiment which he hoped would clarify the situation. Chapter five is Unger's analysis and interpretation of this experiment.

The plan was for *romaji* to be used exclusively in teaching children subjects other than Japanese, and for the performance of these experimental classes to be compared with classes using the usual combination of *kana* and *kanji*. One gets the sense that Unger wanted this little known experiment, which he calls "the most interesting incident in the struggle over script reform" (p. 8) to be the core of his book, and as an advocate of script reform, he was no doubt hoping it would support his cause. The results of the experiment, nearly half a century after its conclusion, however, are uninterpretable largely because the raw data is missing (p. 158), but also because, from what we know of how the experiment was conducted, it is difficult not to concur with Howell V. Calhoun, SCAP's Education Research Officer, who wrote in 1950: "it is hard to find words to describe how completely this project has been bungled" (p. 87).

Although Unger would like to use this experiment to bolster his support for script reform, to his credit he does not shy away from discussion of the experiment's methodological shortcomings. One that he doesn't mention is the experimenters' failure to control for the fact that teachers volunteering to take on the extra work incumbent upon teaching in an experimental program are likely to be, in general, more motivated, hardworking and enthusiastic than the norm. One teacher who volunteered to participate in the *romaji* education experiment, for example, when there were no *romaji* mathematics textbooks available for the class he was to teach, went so far as to have one transcribed at his own expense (p. 94). Further muddying the waters with regard to the quality of the teachers, others reporting on the experiment felt that due to lack of training in *romaji* the volunteer teachers were inferior (p. 105).

Whether they were in fact superior, inferior, or neither one is an empirical question which, unfortunately, cannot be answered today.

That it can't be is, in part, because of the lack of rigor with which the experiment was conducted. *Romaji* might be more effective than *kanji* and *kana*, but because of this lack of rigor, coupled with the lack of raw data, we simply don't, and can't, know.

The hypothesis that the experiment set out to test, "that students who did not have to learn *kanji* as a concomitant part of studying mathematics or the like would make faster progress than students who did" (p. 84), is an interesting one. Those teaching Japanese would, no doubt, still like to know whether using *romaji* would make their jobs easier. It is unlikely, though, that the experiment will be replicated any time soon. The window of opportunity that briefly swung open for language reformers at the end of the war slammed shut again all too soon. As the Cold War intensified the United States found it expedient to rehabilitate the Japanese right. When the conservatives returned to power they brought with them their old resistance to language reform as well as to research which might support it.

Unger's book is most valuable as an object lesson in how, rather than reason or research, it is extralinguistic factors such as politics which ultimately determine the success or failure of language planning and reform. Teachers of Japanese, for example, may see little connection between international relations and the lessons they are planning for next week. If, however, the Cold War hadn't happened, and research on romanization had been allowed to continue, such research may have demonstrated that students whose teachers and texts used *romaji* progressed faster than students whose teachers and texts didn't. If this had been the case, the lessons these hypothetical teachers of Japanese are planning for next week might look rather different than they do.

Literacy and Script Reform in Occupation Japan is, in short, an excellent overview of all the reasons Japanese script should be reformed, and of the reasons it won't be.

Reference

Gottlieb, N. (1995). *Kanji politics: Language policy and Japanese script*. London: Kegan Paul International.

Teaching and Learning in Japan. **Thomas P. Rohlen and Gerald K. Le Tendre (Eds.). Cambridge: Cambridge UP, 1996. 399 pp.**

Reviewed by
Stephen M. Ryan
Osaka Institute of Technology

Popular interest abroad in Japan's education system seems to have waned as the Japanese economy has weakened. This is a great pity as the first wave of journalistic analyses and first-person accounts has given way to a much more considered and enlightening research-based approach to the topic. *Teaching and Learning in Japan* brings together analyses from a surprisingly large number of researchers who have spent long periods of time (years, in some cases) as observers in Japanese schools and other teaching/learning environments.

It is much more, however, than a series of disparate observations. The editors, and many of the authors themselves, have pulled together the separate accounts into a coherent overview of the topic which, nevertheless, manages to eschew stereotypes and present a credibly nuanced analysis of approaches to teaching and learning in this country.

Most of the theory-building occurs in the introductory and concluding essays by Rohlen and Le Tendre but much also depends on their definition of the scope of the book. While the heart of the work is a series of detailed examinations of procedures in and around kindergartens, elementary schools and middle schools, the decision to include chapters on novices in a Zen monastery, new employees at a bank and *noh*-performers, sets the whole discussion in a wider social context which both illuminates and is illuminated by the school-based studies.

Section 1, "Fundamental Approaches," includes essays by G. Victor Sogen Hori and Thomas P. Rohlen. Hori's account of "Teaching and Learning in the Rinzai Zen Monastery" is based on personal experience as a novice. He describes how he learnt to carry out his duties, without any specific instructions, knowing he faced harsh criticism when his performance was less than perfect. Rohlen writes about his experiences as a participant in an induction course for new employees at a bank. His analysis sets in context notorious practices involving physical privations and deliberate social disorientation.

Although, these two essays make sense in their own terms and shed light on the subjects they analyze, it is hard at the time to see how they are meant to fit together, still less how they are relevant to the descriptions of school education which follow, despite the editors' attempts to explain this in the introduction. With hindsight, however, it becomes clear that the "Fundamental Approaches" which these learning experiences share is an over-riding concern for the individual's spiritual development as a member of society. This is a theme which recurs throughout the book.

The second section, "The Emotional Foundations of Early Learning," includes essays by three researchers on separate studies of the education of young children. Catherine Lewis writes about the values which underlie practices in the elementary school classrooms she has observed. As with many of the essays, the comparison with U.S. values and practices is explicit. She asks why it is that, despite the famously longer school year, Japanese children spend so much more time than Americans off-task: holding sports days and class trips, and talking about how to do their best. She concludes that the Japanese ideal is to educate the whole child while the U.S. focus is much narrower.

The comparison continues in Lauren Kautloff's essay which meets head on the allegation that early Japanese education destroys children's individuality. Using illustrations from her observation of pre-school classes, she shows how the efforts of the individual are incorporated into the group, each child being valued for what she brings to the group.

Nancy Sato's essay challenges the whole individual/group dichotomy as unhelpful in seeking to understand Japanese elementary classrooms. Rather than group-oriented, she sees classroom practices as being relationship-oriented, consciously encouraging children to reflect and work on their relationships with each other, with the teacher, with learning materials, and with the subject they are studying.

She also offers a very clear example of a point many of the authors in this collection make: how the surface homogeneity of Japanese schools allows individuals to develop and express themselves. The very fact, she says, that daily routines are standardized and predictable means that teachers can relinquish control of many classroom activities. This both frees the teacher to deal with individual problems and allows each of the students to experience the responsibilities of leadership.

Section 3, "School and Classroom Models," focuses more clearly on the process of instruction in elementary schools. All of the essays here make comparisons with U.S. practices, some of them quantitative as well as qualitative. The picture which emerges from the first three essays will surprise many non-Japanese readers: again and again the classrooms in which inquiry is stimulated, individual opinions nurtured, and responsibility for learning given to the students are the Japanese classrooms rather than the American. The authors here make the point that many successful Japanese practices fulfill ideals espoused by American educators far better than the practices researchers observe in American classrooms.

These conclusions are not especially new (Harold Stevenson in particular has been writing about them for years) but the stereotypes of Japanese education in the U.S. are so fossilized that little impact has been made on the popular imagination. These three essays present, in distilled form, a wealth of research evidence disputing the stereotypes, at least as far as elementary school education is concerned. Once again the impact of the message will probably be minimal among the public at large but we, as educators with a professional interest in Japan, have a responsibility to inform ourselves. Read these three chapters if nothing else.

The fourth essay in this section deals with a mode of education much closer to the stereotype: the Kumon method. Nancy Ukai Russell describes the method and its origins and then draws out its underlying values and beliefs, which she shows to be of a piece with many elements of the mainstream Japanese education system. One of the ways she does this is to analyze how the method has been changed to accommodate American values in the process of adoption in the U.S. This is a very enlightening approach also used by Lois Peak in her later essay on the Suzuki Method of violin instruction.

By this point in the book, I was wondering "What happens to all the lively, inquiring minds and outgoing personalities nurtured by elementary schools before they come into my university classroom?" A partial answer is provided in Section 4, "Path and Guidance," which deals with middle schools. Both of the essays in this section focus on the concept of *shido* (guidance), not just in the sense of verbal advice offered by teachers to students but as a concept that informs every aspect of relations between seniors and juniors in the school.

"While elementary school socialises children to many of the nuances of Japanese life, middle school is the child's introduction to hierarchical organisation and adult patterns of teaching and learning," writes LeTendre (p. 289). Foremost among these patterns is the understanding that learning is a serious business requiring suffering and dedication. Another is that older people (*sempai, sensei* - somebody "born before," veteran teachers) have a clearer understanding of the world than younger ones and are to be respected and obeyed. A strength of this section is the way it brings nuance to these lessons and shows that not everybody follows them wholeheartedly.

I say that the explanation of post-elementary education is only partial because this collection of essays seems incomplete to me. The richness of insights into pre-school and elementary schools is not repeated in the two essays on junior high schools. Very little is said about high schools, despite the near universality of high-school attendance. Rohlen's own (1983) study of high schools makes a useful companion to this volume but suffers by comparison as it presents the perspective of only one researcher. Then there is the resounding silence among social scientists about university education in Japan. When, oh when, is this detailed, multi-perspective approach to be applied to university classrooms?

My plea is heartfelt because, as a university-level language teacher, I see much in the descriptions of elementary school life that I would like to incorporate into my own teaching but before I can contemplate doing so, I need to know much more about how elementary school graduates are socialized by the education process before they enter my classroom.

The final section of the book, "Artistic Pursuits–Old and New," looks at training for *noh*-drama and violin playing. Tom Hare's essay on the concept of "training" in *noh* is rather esoteric but does offer insight into traditional beliefs about appropriate forms of education at different points in the life-cycle. Lois Peak's analysis of the Suzuki Method first explains the basic tenets of the method and then shows how these principles have been changed in establishing the method in the U.S. It might have been better placed alongside the treatment of Kumon.

Despite heavy hints in the introduction, the editors do not include essays on learning in everyday life beyond school-age. The hints suggest that such learning follows patterns established in school education but it would have been helpful to include specific studies showing this to be the case.

This book makes a very important contribution to Western understanding of Japanese education. The approach is thorough, subtle, and convincing. The book should be read by any serious student of Japan and by all who come from abroad and teach here. I hope there will be a second volume, dealing with the education of older children and young adults and with life-long learning.

Reference

Rohlen, T.P. (1983). *Japan's high schools*. Berkeley: University of California Press.

The Self-Directed Teacher: Managing the Learning Process. David Nunan and Clarice Lamb. New York: Cambridge University Press, 1996. 296 pp.

Reviewed by
Ronald M. Honda
Honda Language School

Nunan and Lamb address *The Self-Directed Teacher* to all practicing teachers, embryonic or mature, but primarily to those who are discovering that success in the classroom depends on basing their decisions on the needs, goals, and objectives of the learners and then executing those decisions effectively–essentially, all in-class factors. This, however, is only half of the equation. Nunan and Lamb fail to address the other half, namely managing learning within the external limits placed on each teacher's situation, including institutional flexibility, and the role of outside factors such as parents, peers, and community groups.

The text, its subjects, treatment, educational purpose, and managerial objectives, originated in the authors' many years of discussions and observations in a variety of classrooms and their subsequent observation that autonomous and project-based teaching go hand-in-hand with communicative language teaching. The authors explain how these ideas impact on classrooms worldwide. Their rigorous and labor intensive research work provides a solid foundation from which to begin a project of this nature and magnitude. The general argument can be simply stated: Teachers must effectively utilize their decision-making ability to create suitable environment if students are to learn. In many respects, the authors contend, teachers are unquestionably managers. As such, they should acquire managerial skills so they can create conditions conducive to learning, improve their teaching environment, and adapt to unforeseen situations that arise in the classroom as a result of administrative or executive decisions.

The book is divided into eight distinct and solidly constructed chapters each of which is concluded with supplementary projects and tasks. These supplements appear as extracts providing the teacher with the opportunity to put theory into practice. As a TESOL practitioner himself, this reviewer performed several of the end-of-chapter projects to test the validity of their application in a classroom setting and applied several of the tasks to his own particular teaching situation. While some were not suitable or appropriate in certain settings, most were exceedingly effective. Where a project did not help in one learning situation, it proved useful and suitable in another, illustrating the authors' notion that "in a sense, each learner is an island, and each learner interests a particular classroom even in a slightly different way" (p. 157).

Chapter 1 focuses on the significance of laying a strong foundation of good supervision before the teacher enters the classroom. The authors emphasize the importance of TESOL practitioners establishing a strong basis upon which to expedite classroom management decisions. Most importantly, the efficacy of such decisions depends upon the objectives and desires of the learner. Here the reviewer must raise an objection. The book focuses too tightly on the classroom, neglecting many external factors which influence students and their goals and shapes their responses to classroom activities. For instance, the authors provide no role for parents in the decision-making processes of the teacher. It is a considerable omission since parents play a significant role in their children's education and future. Nobody concerned with the practice of English language teaching, particularly at primary and secondary levels, and particularly in second language acquisition, can ignore the significance of the parents' role in their children's education.

Chapter 2 explores some of the preparation work teachers undertake before entering the classroom. The argument is, for teacher to manage and supervise effectively, their decisions should be structured around a curriculum manifesting an understanding of students' objectives and needs. Nunan and Lamb emphasize the significance of teachers making decisions in the light of students' goals, objectives, and needs, but fail to concretely define these goals and objectives. The authors would strengthen their argument if learner needs and objectives were specified, thus providing teachers with possible solutions and pedagogical guidelines for accommodating such needs. For example, some students may want to concentrate on conversation while others may want to spend most of their time studying syntax and grammar. How does a teacher manage a classroom with such conflicting goals without alienating students?

Chapter 3 takes us into the area of classroom talk, which is essential because it is the way teachers most directly reach learners. We learn the uses of teacher monitoring and why teachers should assess their performance in the classroom. The chapter explores a range of questions, "How much talking do I do?" "To what extent should the teachers employ the students' first language to facilitate their acquisition of the target language?" "Do students get the opportunity to express themselves?" The chapter examines the nature and type of questions teacher ask. The section on managing error is particularly stimulating and teachers in training will find this section quite revealing. However, I question the authors' contention "that learners who have developed skills in identifying their own preferred learning skills and strategies will be more effective language learners" (p. 157), since Nunan and Lamb provide no evidence supporting this claim. For example, I have discovered that I perform well in learning Chinese language in a classroom setting. The question is will I be functional and effective beyond the borders of the classroom?

Chapters 4 and 5 explore time, pacing, classroom monitoring, teacher-learner roles, one-to-one instruction, and self-directed learning, respectively. In an interesting discussion on pacing, the writers start from the two straightforward observations that, since most teachers value constructive use of classroom time, particularly when there is a set amount of material to cover, they must decide how long activities should last and need to be aware of the time available. Nunan and Lamb then note the deeper requirement, "Before managing our time, we need to find out what we are actually doing with it" (p. 126). The discussion of how teachers make effective decisions with the amount of resources at their disposal is particularly interesting.

This section also discusses the constructive use of time in response to cultural factors and behavioral problems teacher must encounter daily, such as cultural differences over roles and rules between teacher and students. Most of all, Nunan and Lamb stress, teachers must be firm, for as Harmer (1991, p. 249) points out, "one way of avoiding most disruptive behavior (though not all) is by making sure that all your students of whatever age know 'where you stand'." This chapter also fails to include the parents in the equation as an important element in resolving discipline problems of students. A section on the role of parents in behavior problems would develop the discussion of classroom discipline. Nevertheless, the classroom snaps (pp. 135-136) are particularly useful.

With increasing numbers of ESL teachers exploring the field of language brokering on their own, the brief section on one-to-one instruction needs more extensive treatment. The number of teachers in Japan giving private lessons has grown over the past five years. A similar state of affairs exists in China and other Asian countries where there is a great demand for English teachers. Nevertheless, these chapters are quite informative and enlightening, and this reviewer particularly enjoyed the interview skit with teacher and student on pages 150-152.

In chapters 6 and 7 Nunan and Lamb discuss managing resources, motivation, attitude, and aptitude. In general chapter 6 notes that commercially produced teaching materials, if used constructively and with the needs and objectives of learners in mind, give the teacher flexibility to achieve classroom goals. To achieve this, the writers emphasize the establishment of criteria to follow when selecting course materials. We learn that the effective use of such resources is crucial to the successful management of the classroom.

The authors also discuss the roles of motivation, attitude, and aptitude in second language acquisition. This section stresses the teacher's responsibility to find creative ways of enhancing students' motivation and developing their attitude toward the learning process (in addition to whatever aptitude they bring to the language). All this assumes, of course, that teachers have sufficient latitude and independence in the classroom to allow for such flexibility. A teacher without such latitude is at a serious disadvantage and those who operate within a framework of strict limitations often find creativity stifled.

The final chapter discusses at length the importance of self-monitoring and evaluation among teachers in a critically but constructive manner. Evaluation among teachers, whether formal or informal, serves a crucial role in each teacher's decision-making process. Not only does such assessment save time in planning, it also plays a significant role in the learner's progress.

The book has three main uses. Administrators, directors, and principals of schools can use it as they try to better understand and accommodate their teaching staff. Teachers and teachers-in-training will want its insights to help them evaluate and assess their effectiveness as educators and to refine and develop their style and managerial and supervisory acumen in the classroom. Finally, the book's focus on self-direction and its emphasis on teachers' awareness of students' needs makes clear the importance of autonomy in the classroom, for teachers and students. The book appeals to such a broad audience–teachers, teachers-in-training, administrators, supervisors, and principles of schools–by not imposing a particular pedagogical approach to teaching and by encouraging teachers and administrators to seek alternative ways of dealing with the daily managerial problems of teaching by always keeping students' needs in mind.

Undoubtedly thousands of books have been written on management. Just take a look in the management section of any bookstore. This book, however, goes beyond them. It takes a different view, focusing primarily on classroom management from a teacher-student perspective and shows teachers how to use management skills to enhance their effectiveness as teachers and to stimulate learner involvement. This makes *The Self-Directed Teacher* unique. Most books on management concern themselves with financial printouts and profit and loss aspects of business; the profit factor, if you will. Rather this volume deals with the human aspects of management and devotes special attention to the needs and objectives of the learner as customer in institutions of learning. The authors have, therefore, written a management text from the perspective of student and teacher. In this sense the book is an invaluable resource and is a major contribution to the development of the teacher-student approach to management with one significant omission: conflict management.

The subject of conflict management is crucial to a teacher's management skills and should have been explored in more detail. Situations always arise where there is a breakdown in communication, between teachers and management, and between teachers and students. A chapter on how to resolve such conflicts would have been helpful. This minor oversight does not, however, take away from the book's effectiveness.

Nunan and Lamb, however, have still managed to give us in *The Self-Directed Teacher* an even document of useful research, thought provoking issues, and above all, one that teachers will find of practical use.

Reference

Harmer, J. (1991). *The practice of language teaching*. London: Longman.

Phonology in English Language Teaching: An International Approach. **Martha C. Pennington. (Applied Linguistics and Language Study series). London and New York: Longman, 1996. xviii + 282 pp.**

Reviewed by
Ron Grove
Mejiro University

Martha Pennington's background prepared her well for making the principles of phonology accessible to those engaged in or training for ELT. Born in Florida, she moved to New Hampshire at age 13 and still recalls the accent shock she and her new speech community experienced. She studied and taught applied linguistics in Pennsylvania, California, and Hawaii, later also at Nagoya Gakuin University and Temple University Japan; she is now Professor of English at the City University of Hong Kong (pp. xiv-xv). Her perspective is "variationist, accent-neutral and international" (p. xvi), both politically correct and practical for JALT members.

Although Pennington does not spare technical terms and writes very dense prose that cannot be read casually, she also explains things clearly, usually with helpful examples. For instance, if "a combination of retroflex and labialized articulation" fails to evoke anything in your mind's ear, the example offered that this speech quality "made the American actor, Jimmy Stewart, seem ingenuous" (p. 160) makes the effect easier to imagine or test for yourself.

Pennington's initial advice to readers is that her book is a multidisciplinary "comprehensive introduction to English phonology" aimed primarily at teachers of English as a second or foreign language (p. xvi), and it is exactly that. There are six chapters: 1) Introduction to phonology in language teaching; 2) Consonants; 3) Vowels; 4) Prosody (e.g., stress, intonation); 5) Phonology and orthography, and; 6) Pronunciation in the language curriculum. The main text of each is followed by extensive "activities" by which readers can check or apply their understanding of the material. The first five chapters also include "teaching ideas." The few that I tried with my college classes worked well and were enjoyable. There are also three appendices: A) Hierarchical analysis of student pronunciation; B) Pedagogical classification of pronunciation errors and problems, and; C) Sample unit plan for teaching the /r/-/l/ distinction.

This last is not the only place where the author's East Asian experience may have contributed to the development of her ideas. Her warning about the difficulty of distinguishing "foreign" from "native" accents (pp. 6-7) and her citation of Nigerian E. Adegbija to the effect that native speakers are not necessarily the best teachers of a target foreign language being learned for community-internal purposes (p. 240) could usefully correct the "native-speaker [of English]" mystique current in this country, particularly if accompanied by a realistic and critical examination of the goals of ELT here. Sometimes her awareness of English phonological issues relevant to Japan comes out in amusing ways. Teaching ideas involving the "/w/, /v/ , /f/, /h/ contrast (e.g. for Japense [sic] students)" include the following awareness activity. When a Japanese student says "manfood," what does s/he mean? (1) food made for humans (cf. dogfood); (2) the condition of being a man (manhood); (3) part of a conditional statement about a man (man would); (4) part of a relative clause statement about a man (man who'd); (5) any or all of the above (p. 84).

In an otherwise exemplary textbook, there are a few spots in need of more careful editing or proofreading. Some errors, like the "Japense" students, mentioned above, are obvious and not problematical, but not all. The prose is so precise and economical that it took me, at least, quite a bit of thought to satisfy myself that the passage below, rather than my understanding of it, contained a serious error.

> A tendency for complementarity in length of the consonant and vowel in syllables made up of Vowel+Consonant (VC), and to a lesser extent in CV syllables, has been found for English. According to this tendency, a durationally long consonant is preceded (or followed) by a short vowel, and a durationally short consonant is preceded (or followed) by a short vowel (p. 100).

The sentences following this passage help clarify that the last two words should in fact be "long vowel."

In another activity in which intonation alone is to be a clue whether a sentence is a statement or a question, all pairs are verbally identical, except for one: "There are three yellow ones. There are only three yellow ones?" (p. 170).

In some cases, the economy with which a situation is introduced makes it hard to interpret, e.g., the following question for an activity: "Why does a cold cause /m/ to be denasalized, with [b] substituting for /m/?" (p. 72). A more natural way to bring up the topic of what it sounds like to talk through a cold ("When you have a cold, why do you tend to denasalize /m/, replacing it with [b]?") would have helped. The sudden introduction of "a cold," unaccompanied by usual collocations, is disorienting.

Less important, but equally bizarre is the stylistic decision to refer to certain quoted texts, printed in the same font as the main text, as numbered "figures" (e.g., "figure 5.1," p. 186, is a poem). The respect for ordinary English usage evident throughout Pennington's treatment of the sound system should be extended to the use of words as well.

This book is exactly what it says it is. It systematically explains the scientific phonology of English as it relates to teaching and learning that language. Although features of other sound systems are sometimes mentioned, this is not an all-purpose phonology text. Although some persons other than those concerned with ELT may find it useful, it was not written with anyone else in mind. I would recommend it as a very useful coursebook without reservation to anyone engaged in training teachers of English as a second or foreign language, as well as to those who wish to acquire the equivalent of a good graduate-level introduction to English phonology through self-study.

Vocabulary, Semantics, and Language Education. **Evelyn Hatch and Cheryl Brown. Cambridge. Cambridge University Press, 1995. 468 pp.**

Reviewed by
Mark O'Neil
ALC Education

This book, from the Cambridge Language Teaching Library, a companion volume to Discourse and Language Teaching: *Vocabulary, Semantics, and Language Education,* according to the back cover, focuses on "the exploration of semantic and lexical theory and the practical application of this theory to language teaching and language learning." The book does not provide the language teacher with any practical techniques for vocabulary teaching, but rather attempts to draw the reader to an understanding of why a language learner may or may not successfully acquire lexical processes.

The book has five parts, "Semantics," "Lexicon," "Lexical cases and morphology," "Vocabulary choice and discourse use," and "Vocabulary learning and vocabulary teaching." Each chapter contains a variety of activities for practice and review as well as suggestions for further research.

The first part, "Semantics," begins with a description of semantic feature analysis and introduces the different motivations of psychologists and linguists for describing language. The presentation of a variety of models such as componential analysis, core meanings and prototype theory, and relational models discusses the strengths and limitations of each model. Linguists, on the one hand, are searching for a description of semantics that fits into a complete description of language. In the search for such a description they break language down into increasingly smaller and smaller components in order to classify not just words but also the various separate and different meanings of those words. On the other hand, psychologists are concerned primarily with the relationship between perception and language. Part I also describes the way we use language in a less than literal way.

Part II, "Lexicon," describes the ways the vocabulary of a language increases, including borrowing from other languages, coinage, compounding, and clipping. Hatch and Brown discuss these processes along with their implications for language learners.

The third part, "Lexical cases and morphology," examines traditional methods of classifying words into parts of speech and presents some difficulties arising from such classifications. The authors follow this by a look at word formation processes with derivational morphology, including a discussion of the processes that can lead to learner error. There is also a chapter on inflectional morphology, a grammar-centered description, included here as part of the lexical system.

Part IV, "Vocabulary choice and discourse use," describes of how lexical choice is affected by register, style, gender and other social factors. It also addresses the way that groups or individuals express their uniqueness, biases and preferences through word choice.

"Vocabulary learning and vocabulary teaching," the final part, takes the understanding gained of semantics and the lexicon in the first four parts and applies it to language learning and language teaching. The first chapter in this part describes learners' strategies for vocabulary acquisition , "five essential steps in vocabulary learning," and the ways in which they may compensate for words they do not know. Then comes a chapter on the strategies used by teachers and material writers to help learners understand and learn vocabulary.

The book is an introduction of the study of semantics and the lexicon drawing on a wide variety of research and supported by a large number of examples. It states in the preface that this book was originally prepared as part of an introductory course in linguistics at the University of California, and it does have a "book of the course" feel to it. There is an extensive list of references to which the authors direct those who wish to gain a more detailed understanding of the field.

While the practice exercises, which are included in every chapter, are a very welcome and thought-provoking element of the book, some of the suggested activities or ideas for research could be quite time-consuming and may well be more practically approached by a group of linguistics students or trainee teachers within a cooperative framework rather than by an individual language teacher working alone.

Vocabulary, Semantics and Language Education in many ways is two books in one, either or both of which might interest *JALT Journal* readers. The first four parts are aimed at the student of linguistics; the final part focuses on pedagogy and the needs of practicing language teachers. Those looking for a solid presentation of the theoretical and linguistic bases of vocabulary and semantics will find the first part most valuable. Language teachers looking for pedagogical ideas for facilitating vocabulary learning or wishing to read about what strategies language students employ to learn vocabulary could simply skip straight to part five. In either case, Hatch and Brown have written a book of general interest and practical importance.

Disorders of Discourse. Ruth Wodak. London: Longman, 1996, 200 pp.

Reviewed by
Sandra Ishikawa
Osaka

The title of Ruth Wodak's book, *Disorders of Discourse*, is ambiguous. It could refer to several possible types of disorder causing communication problems. It might be about special cases like mental health problems, or about brain damage. It might be about normal people or non-native speakers having communication problems in ordinary conversation. The disorders in the title are those of ordinary discourse, but of a more subtle kind, those occurring in organizations. Such disordered discourse may involve people in organizations in relation to people outside, in client, patient or other subordinate roles. It may instead involve individuals or groups inside organizations who differ in the amount of power they have. This book presents a method of analyzing the use of power to control, subvert, prevent, or merely fail to promote discourse. These are the disorders of the title, disorders sometimes so subtle that the people involved may be unaware of any discourse failure. Clearly, this is a method with a mission.

Wodak, who works in Vienna, describes her book as presenting a method of analyzing discourse, which has developed over many years of research, from the 1970s through to the present (p. 3). The research provides the examples which demonstrate how the approach works. In the acknowledgments, Wodak points out that much of the research has been collaborative, involving both her colleagues and students (pp. ix-x). The first chapter, the introduction, is the densest. Its 34 pages lay out the background, definitions, rationale, method, and the impact of the research. Each of the next four chapters presents a research project demonstrating the approach. The sixth and last chapter, only eleven pages long, is both a summary and the conclusion. A bibliography and an index follow. Wodak does not mention her target audience, but the back cover claims the book is "relevant for students and academics in linguistics, sociology, psychology and education."

Chapter one begins by presenting sample communications which are not, or not easily, understood by their targets. The samples come from a government tax bureau, a doctor and a news broadcast, and represent "frame conflicts." A frame conflict arises when "worlds of knowledge and interests collide with one another, and those who possess linguistic as well as institutional power invariably prevail" (p. 2). Wodak then presents evidence that people inside organizations also fail to communicate successfully with each other. These two types of organizational discourse, internal and external, are the subject matter for analysis. "Internal" refers to employees of the organization, of whatever rank, and "external" refers to the clients, whether they are patients, pupils, parents, or radio listeners. This area of study, which she calls "discourse sociolinguistics," looks at both discourse production and discourse comprehension (p. 3). Ethnographic techniques are used to get an insider perspective on the organizations under study. Tests and interviews are used to get the outsider perspective (p. 3). The information gained is used to suggest changes. The approach can also be used to analyze the effect of making such changes.

Continuing with a brief historical overview, Wodak describes her method of analysis as related to several sociolinguistic trends, especially (a) analyzing "text in context," (b) interdisciplinary studies, and (c) political concerns (pp. 5-6). She then presents a number of definitions, and devotes several pages to critical discourse analysis, which has strongly influenced her work. She next discusses methodology, and the viewpoints of several researchers. She explains her decision to use both micro- and macro-analysis and multiple methods. The final sections deal with the implications of such research from the viewpoint of society. This first chapter concludes:

> As discourse sociolinguists, we provide instruments for a less authoritarian discourse. And such instruments may, but do not have to, lead to emancipation. Thus, the results of our studies are important in many ways. First, they make transparent inequality and domination. Secondly, they enable us to propose possibilities of change. And, thirdly, they show the limits of possible emancipation through new patterns of discourse alone. (p. 32)

Having laid out this background, Wodak devotes the next four chapters to presenting four quite different studies demonstrating the discourse sociolinguistic approach. These chapters are easier reading than the first. Chapter two describes disorders of discourse in hospitals. Doctor-patient, or insider-outsider, discourse is taken up first. An overview describes the setting and research methods, as well as the categories of analysis. Three case studies follow. In the first two, an inexperienced patient and an experienced patient are handled quite differently, indicating that a patient's familiarity with medical matters, especially jargon, affect how the doctor behaves toward the patient. The third case study reports on a patient whose time with the doctor is repeatedly interrupted by outside problems. This case leads into a discussion of myths which hospital personal perpetuate and cooperate in. These myths are discussed separately. Analysis shows that the myths may be intended to increase harmony among insiders, particularly doctors and nurses, but they actually contribute to some of the problems, both internal and external. The chapter closes with a comparison of the type of information obtained by this approach and the type obtained by more common methods of discourse analysis.

Chapter three describes disorders of discourse in schools. These again are both internal (among employees—heads and teachers of various ranks), and external (between these heads and teachers, and the clients—pupils and parents). This research emerged from a 1985 Austrian law intended to increase democratization by requiring "school partnership" among the various interested parties. Chapter six informs us that the research was in fact commissioned by the Ministry of Education to measure the effect of this law (p. 173). Three types of school are examined with several techniques, including quantitative and qualitative measures. Stages and categories of analysis are presented. Several meetings are presented, with transcripts and analyses, showing how control and manipulation occur. The chapter concludes by pointing out that the law has not led to greater democratization, but instead, "power structures have been reproduced more subtly and have thereby become even more difficult to oppose (p. 98).

The next chapter, the fourth, looks at news broadcasts. News stories differ from ordinary stories in ways that make news stories difficult to understand and to update as further information becomes available. There is a discussion of what comprehension is and of various types of schemata. and then a model of cyclical comprehension. Two experimental studies are reported here. One involved 277 students, and the other fifty adults (18 to 73 years old). Text manipulation showed that changes in the texts increased understanding of the news, but that greater increases were shown by those who were already better informed. Comprehension increased, but social class and gender differences increased more.

Chapter five reports a three-year study of the Vienna Crisis Intervention Centre, and focuses on communication among therapists and patient treatment groups. She discusses the ethics of research under such circumstances, and describes the setting, the study and the hypotheses. She introduces a model of therapeutic communication, with three levels of "meaning:" (a) colloquial; (b) group-specific; and (c) private (pp. 143-146). She analyzes the text of the interactions in terms of the meaning levels, and of "moves," although she does not use the term. "Moves" are steps which typically occur in specific sequences and forms in various genres. [See, for example, Swales (1990) for a discussion of moves in research papers, and Connor (1996) for a discussion of moves, including Swales'.] The discourse genre in the present study involves problem presentation through such moves as opening focus, scene, narration, circumstance, and closing focus. The text analysis uses the levels of meaning in combination with the moves to examine the progress of therapeutic discourse in therapist-patient group discussions.

The eleven pages in Chapter six comprise a summary each of the chapters presenting research, chapters two through five. Each summary concludes with a discussion of what this research approach can do in that situation, what problems there are, and what social changes this method can and cannot aid. By including a summary for each research type in this chapter, a broad view of the method is obtained. The reader can see how the approach is adapted to different goals and different settings.

This book should be considered from several perspectives. From the first perspective, it introduces an approach to analyzing "text in context" using a wide range of methods, with a view toward initiating or measuring social change. In this it is very convincing. It is difficult to imagine any other method producing such complete and well-documented evidence of the relationships between interaction, participants, purposes and power. This combination of ethnography, text analysis, and other techniques results in impressive breadth and depth of understanding. Although the author does not make claims outside the specific area of each study, anyone familiar with a large variety of meetings in Japan will recognize the school research (chapter three) as describing such meetings very well. Shortly after reading the chapter, I attended a meeting of condominium owners which was run by the board of directors in exactly the way, and with exactly the intentions, described in the chapter. The annual company stockholder meetings are another example of this type of control.

Although the use of analyses such as these for political purposes may be unfamiliar and even surprising to most readers, they may come to appreciate how well such work can explore relationships within organizations and between them and their clients. Such research can not only look for ways to change relationships, but can also evaluate the changes to see if they have met the intended goals, which these studies show they often have not. The first perspective, then, is introducing the reader to the potentials of this type of research approach.

The second perspective follows from the first. Few of the readers of *JALT Journal* will have the resources and the power to undertake some of the studies described here. Nor would language teachers want to carry them out. The radio news studies and smaller-scale classroom studies modeled on the other studies are perhaps the most likely applications to interest language classroom researchers. The book offers more for the researcher than for the teacher in the language classroom.

A third perspective is that of a member of society who is involved in the kinds of situation analyzed in the book in our daily lives (excepting the group therapy sessions, of course). We interact with various organizations, require medical attention, attend meetings in various capacities, and listen to news broadcasts. The analyses here can increase our awareness of how we manipulate people or are manipulated by them. As Wodak points out, the first step to initiating change is to become aware of what is happening.

As a whole, the book is broadly informative. The admittedly political intentions may put off some readers, but they are not offensive, and readers should try to look past them to see what is intended. They may find themselves persuaded, as I did, that there is a great deal of need for such analyses in many situations involving power-holders and their subordinates or clients. There is also a great potential for multidimensional approaches like Wodak's.

Criticisms of the book are minor. The paper used is permeable, so highlighting goes right through the page. There are a few instances where the language is not quite native. The tables on pages 118 and 119 report the significance levels of statistical procedures as "*a*" when they should surely be "*p*." The chapters reporting research , two through five, do not have closing sections. The summaries, conclusions, and implications for these are found in Chapter six. Chapter six, although brief, concludes the book very well, but the reader might like to know this information at an earlier point. On the whole this is an interesting book. It was not what I expected, given the title, but it is well worth reading.

References

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing*. Cambridge: Cambridge University Press.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Books to Review

Reviewers are being sought for the following texts. Contact the Reviews Editor, see Guidelines, for further information.

Baynham, M. (1995). *Literary practices: Investigating literacy in social contexts.*

Durand, J., & Katamba, F. (1995). *Frontiers of phonology: Atoms, structures, derivations.*

Ferris, C. (1993). *The meaning of syntax: A study in the adjectives of English.*

Jacobs, R.A. (1995). *English syntax: A grammar for English language professionals.*

Kunnan, A.J. (1995). *Test taker characteristics and test performance: A structural modeling approach.*

Ridout, R.M. (Ed.) (1996). *The Newbury House dictionary of American English: An essential reference for learners of American English and culture.*

Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quanitative and qualitative analyses.*

Spada, N., & Frolich, M. (1995). *COLT (communicative orientation of language teaching) observation scheme: Coding, conventions and applications.*

Street, B.V. (1995). *Social literacies: Critical approaches to literacy in development, ethnography and education.*

Information for Contributors

Contributions must conform to the *JALT Journal* Editorial Policy and Guidelines.

Editorial Policy

JALT Journal, the research journal of the Japan Association for Language Teaching (*Zenkoku Gogaku Kyoiku Gakkai*), welcomes practical and theoretical articles concerned with foreign language teaching and learning in Japanese, Asian, and international contexts. It welcomes well written contributions which provide readers with a link between theoretical and practical issues, especially those addressing current concerns in pedagogy, methods, and applied linguistics. Areas of specific interest are:

1. curriculum and teaching methods
2. classroom centered research
3. cross-cultural studies
4. teacher training
5. language learning and acquisition
6. overviews of research and practice in related fields

The editors encourage submission of full-length articles, short research reports, book and media reviews, essays on trends in language education, reports of pedagogical techniques which are thoroughly framed in theory and supported by descriptive or empirical data, and comments on previously published *JALT Journal* articles. Occasionally, *JALT Journal* will issue a Call for Papers for theme-based issues. Articles should be written with a general audience of language educators in mind, with statistical techniques and unfamiliar terms clearly explained or defined.

Guidelines

Style

JALT Journal uses the *Publication Manual of the American Psychological Association*, 4th edition (available from the Order Department, A.P.A., 1200 17th St. N.W., Washington DC). Consult recent copies of *JALT Journal* or *TESOL Quarterly* for examples of documentation and references. Give the page numbers of cited works in both the text and references.

Format

No more than 20 pages (approximately 6000 words) for full-length articles, including reference list, typed and double-spaced on A4 or 8.5"x11" paper. *Research Forum* submissions should be no more than 10 pages. *Perspectives* submissions should be no more than 15 pages. Authors must supply camera-ready diagrams or figures (if any) before final publication. The author's name and references that identify the author should appear only on the cover sheet.

Materials to be submitted

Three (3) copies of the manuscript, no author reference
Cover sheet with title, running head title (one to four words), and author name(s)
Abstract (no more than 150 words)
Japanese translation of title and abstract if possible (less than 400 *ji*)
Biographical sketch(es) (no more than 50 words)
Authors of accepted manuscripts must supply a disk copy (RTF or ASCII).

Evaluation Procedures

All manuscripts are first reviewed by the editorial staff to insure they comply with *JALT Journal* guidelines. Those considered for publication are subject to blind review by at least two readers, with special attention given to the *JALT Journal* Editorial Policy, the significance and originality of the submission, and the use of appropriate research design and methodology. Evaluation is usually completed within three months.

Restrictions

Papers submitted to *JALT Journal* must not have been previously published, nor should they be under consideration for publication elsewhere. *JALT Journal* has First World Publication Rights, as defined by International Copyright Conventions, for all manuscripts published. We regret that manuscripts or computer disks cannot be returned.

Address for Manuscripts and Inquiries

Sandra Fotos, *JALT Journal* Editor (incoming)
Senshu University, 2-1-1 Higashi Mita
Tama-ku, Kawasaki, Kanagawa 214 Japan

Submissions to *Perspectives*

The editors invite reports of pedagogical techniques which are thoroughly framed in theory and supported by descriptive or empirical data to the *Perspectives* section. Send submissions to:

Nicholas O. Jungheim, *JALT Journal* Associate Editor (incoming)
Aoyama Gakuin University, Faculty of Law
4-21-25 Shibuya, Shibuya-ku, Tokyo 150 Japan

Reviews and Address for Reviews

The editors invite reviews of books, tests, teaching systems, and other substantial publications in the field of language education. A list of publications which have been sent to JALT for review is published monthly in *The Language Teacher*. Please send requests for books, materials, reviews guidelines, and queries to submit reviews to:

Thomas Hardy, *JALT Journal* Reviews Editor
Tamagawa University, 6-1-1 Tamagawa Gakuen
Machida-shi, Tokyo 194 Japan

Japanese-Language Manuscripts

The *JALT Journal* welcomes contributions on teaching Japanese as a second language (JSL). Submissions must conform to the above Guidelines. Authors are requested to proved a detailed abstract in English, 500-750 words in length. Refer to the Japanese-language Guidelines for details. Send all Japanese-language manuscripts to:

Naoko Aoki, *JALT Journal* Japanese-language Editor
Shizuoka University, Department of Education
836 Oya, Shizuoka 422 Japan

Address for Inquiries about Subscriptions or Advertising

JALT Central Office
Urban Edge Bldg. 5F
1-37-9 Taito, Taito-ku, Tokyo 110 Japan
Tel: 03-3837-1630; Fax: 03-3837-1631
(From outside Japan: Tel: 81-3-3837-1630; Fax: 81-3-3837-1631)

JALT Journal 日本語論文投稿要領

JALT Journal は、日本語教育に関する日本語論文の投稿を受け付けます。第二言語教育／学習研究の文脈を踏まえ、しっかりした研究計画に基づいている、実践への応用の可能性を示した理論的・実証的研究や理論的裏づけを持った実践報告などで、教育、教授法、応用言語学などの今日的問題を扱ったものを歓迎します。

文体は一般的な学術論文のスタイルを使ってください。章だてのしかたや参考文献のデータの書き方などは、Publication Manual of the American Psychological Association (4th. ed.) の定める方式にできるだけ近い形にしてください。ご不明の場合は、JALT Journal の英語論文を参考にするか、日本語編集者までお問い合わせください。また、JALT Journal の読者は現場の教師が主ですから、特殊な専門用語や統計的手法は、わかりやすいように定義したり説明したりしてください。

原稿の長さは、参考文献リストも含め400字づめ原稿用紙30枚以内です。A4の用紙に横書きで、1行40字、1ページ30行で印刷してください。手書きの原稿は受け付けられません。図表を入れる場合は、JALT Journal のページのサイズに合わせて作成してください(縮小コピー可)。執筆者の名前や所属機関名は、原稿には書かないでください。ご提出いただくものは、以下の通りです。

- 原稿3部
- 執筆者の名前と所属機関名を書いた表紙
- 400字以内の和文要旨
- 英文のタイトルと、500～750語の英文要旨
  英文要旨が用意できない場合は、400字づめ原稿用紙3枚程度の詳細な和文要旨をご提出ください。
- 100字以内の執筆者略歴
- 可能ならば、Macintosh のテキスト形式で保存したファイル

査読の方法は、編集委員会で投稿要領にあっているかどうかを確認したあと、少なくとも二人の査読者が査読を行います。査読者には執筆者の名前は知らされません。査読の過程では特に、原稿がJALT Journal の目的にかなっているか、言語教育にとって意味があるか、独創性はあるか、研究計画と方法論が適切かが評価されます。査読は通常二か月以内に終了します。

JALT Journal に投稿する原稿は、すでに出版されているものや他の学術雑誌に投稿中のものは避けてください。JALT Journal は、そこに掲載されるすべての論文に関して国際著作権協定による世界初版権をもちます。

お送りいただいた原稿は返却できません。

投稿原稿の送り先とお問い合わせは以下にお願いします。

〒422　静岡市大谷836　静岡大学教育学部
JALT Journal 日本語編集者　青木直子

JALT Journal 第19巻　第2号

1997年11月　7日　印刷
1997年11月10日　発行

編集人　タマラ・スウェンソン
発行人　ジーン・ヴァン・トロイヤー
発行所　全国語学教育学会事務局
〒１１０東京都台東区台東１－３７－９アーバンエジビル５Ｆ
TEL (03)3837-1630; FAX (03)3837-1631
印刷所　コーシンシャ株式会社
〒530 大阪市北区天満2-13-3 TEL (06)351-8795

BEST COPY AVAILABLE

Phrasal verbs are a major feature of the English language. Knowing how to use them, or what they really mean, can be a real problem.

The Cambridge International Dictionary of Phrasal Verbs gives easy to find explanations that are also easy to understand.

Clear
Detailed
Extensive

wise up *British slang: Cambridge will guide you to the real meaning*

Cambridge University Press
c/o UPS, Kenkyusha Bldg.
2–9 Kanda Surugadai
Chiyodu-ku, Tokyo 101
Tel: (81)-3-3295-5875 Fax: (81)-3-3219-7182

CAMBRIDGE
UNIVERSITY PRESS

Enough said? It's one thing knowing **how to write words in English**, quite another thing knowing how to say them.

The English Pronouncing Dictionary shows users exactly how to say more than 80,000 words in English. This is the authoritative guide to pronunciation.

enough ɪ'nʌf *Cambridge will guide you to the real pronunciation.*

BEST COPY AVAILABL

JALT Journal

Information about *JALT Journal* can be found at:
www.cc.aoyama.ac.jp/htupdate/aoyama/jjweb/jj-index.html

On-Line

全国語学教育学会
JALT
JOURNAL

JALT Conference Calendar

JALT 98 - Omiya, November 20 - 23

Focus on the Classroom: Interpretations

CALL for Papers Deadline: Feb. 1, 1998
Contact: JALT Central Office, Urban Edge Bldg. 5F,
1-37-9 Taito-ku, Tokyo 110, Japan

Other Conferences in Japan

PacSLRF 98, Tokyo, March 26 - 29

Contact: PacSLRF 98, Aoyama Gakuin University,
Dept. of English, 4-4-25 Shibuya, Shibuya -ku, Tokyo 150, Japan